JMIR mHealth and uHealth

Mobile and tablet apps, ubiquitous and pervasive computing, wearable computing and domotics for health
Volume 10 (2022), Issue 1 ISSN: 2291-5222

Contents

Reviews

Human-Centered Design of Mobile Health Apps for Older Adults: Systematic Review and Narrative Synthesis (e29512)
Zethapong Nimmanterdwong, Suchaya Boonviriya, Pisit Tangkijvanich. 4

Promoting Physical Activity and Weight Loss With mHealth Interventions Among Workers: Systematic Review and Meta-analysis of Randomized Controlled Trials (e30682)
Jiyeon Jung, Inhae Cho. 52

Use of Mobile Apps for Self-care in People With Parkinson Disease: Systematic Review (e33944)
JuHee Lee, Insun Yeom, Misook Chung, Yielin Kim, Subin Yoo, Eunyoung Kim. 70

The Impact of Wearable Technologies in Health Research: Scoping Review (e34384)
Sophie Huhn, Miriam Axt, Hanns-Christian Gunga, Martina Maggioni, Stephen Munga, David Obor, Ali Sié, Valentin Boudo, Aditi Bunker, Rainer Sauerborn, Till Bärnighausen, Sandra Barteit. 88

The Influence of Design and Implementation Characteristics on the Use of Maternal Mobile Health Interventions in Kenya: Systematic Literature Review (e22093)
Karen Sowon, Priscilla Maliwichi, Wallace Chigona. 117

Impact of Smartphone App–Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials (e29621)
Maria Serrano-Ripoll, Rocío Zamanillo-Campos, Maria Fiol-DeRoque, Adoración Castro, Ignacio Ricci-Cabello. 133

Efficacy, Effectiveness, and Quality of Resilience-Building Mobile Health Apps for Military, Veteran, and Public Safety Personnel Populations: Scoping Literature Review and App Evaluation (e26453)
Melissa Voth, Shannon Chisholm, Hannah Sollid, Chelsea Jones, Lorraine Smith-MacDonald, Suzette Brémault-Phillips. 327

Original Papers

mHealth Solutions for Perinatal Mental Health: Scoping Review and Appraisal Following the mHealth Index and Navigation Database Framework (e30724)
Benedetta Spadaro, Nayra Martin-Key, Erin Funnell, Sabine Bahn. 25

XSL•FO
RenderX

Evaluating Evidence-Based Content, Features of Exercise Instruction, and Expert Involvement in Physical Activity Apps for Pregnant Women: Systematic Search and Content Analysis (e31607)
Melanie Hayman, Kristie-Lee Alfrey, Kim Waters, Summer Cannon, Gregore Mielke, Shelley Keating, Gabriela Mena, Michelle Mottola, Kelly Evenson, Margie Davenport, S Barlow, Emily Budzynski-Seymour, Natalie Comardelle, Madison Dickey, Cheryce Harrison, Maryam Kebbe, Trine Moholdt, Lisa Moran, Taniya Nagpal, Stephanie Schoeppe, Stephanie Alley, Wendy Brown, Susan Williams, Lisa Vincze. 37

Knowledge and Expectations of Hearing Aid Apps Among Smartphone Users and Hearing Professionals: Cross-sectional Survey (e27809)
Jae Han, Yong-Ho Park, Jae-Jun Song, Il Moon, Woojoo Lee, Yoonjoong Kim, Young Cho, Jae-Hyun Seo, Moo Park. 148

Applying an Extended UTAUT2 Model to Explain User Acceptance of Lifestyle and Therapy Mobile Health Apps: Survey Study (e27095)
Eva-Maria Schomakers, Chantal Lidynia, Luisa Vervier, André Calero Valdez, Martina Ziefle. 160

Assessing Elderly User Preference for Telehealth Solutions in China: Exploratory Quantitative Study (e27272)
Nuoya Chen, Pengqi Liu. 176

Dose–Response Effects of Patient Engagement on Health Outcomes in an mHealth Intervention: Secondary Analysis of a Randomized Controlled Trial (e25586)
Yiran Li, Yan Guo, Y Hong, Yu Zeng, Aliza Monroe-Wise, Chengbo Zeng, Mengting Zhu, Hanxi Zhang, Jiaying Qiao, Zhimeng Xu, Weiping Cai, Linghua Li, Cong Liu. 196

Effect of an Integrative Mobile Health Intervention in Patients With Hypertension and Diabetes: Crossover Study (e27192)
Sang Oh, Kyoung-Kon Kim, Sung Kim, Su Park, Sangshin Park. 208

The Association Between Home Stay and Symptom Severity in Major Depressive Disorder: Preliminary Findings From a Multicenter Observational Study Using Geolocation Data From Smartphones (e28095)
Petroula Laiou, Dzmitry Kaliukhovich, Amos Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Shaoxiong Sun, Yuezhou Zhang, Faith Matcham, Alina Ivan, Grace Lavelle, Sara Siddi, Femke Lamers, Brenda Penninx, Josep Haro, Peter Annas, Nicholas Cummins, Srinivasan Vairavan, Nikolay Manyakov, Vaibhav Narayan, Richard Dobson, Matthew Hotopf, RADAR-CNS. 222

A Mobile App to Increase Fruit and Vegetable Acceptance Among Finnish and Polish Preschoolers: Randomized Trial (e30352)
Henna Vepsäläinen, Essi Skaffari, Katarzyna Wojtkowska, Julia Barli ska, Satu Kinnunen, Riikka Makkonen, Maria Heikkilä, Mikko Lehtovirta, Carola Ray, Eira Suhonen, Jaakko Nevalainen, Nina Sajaniemi, Maijaliisa Erkkola. 235

Investigating When, Which, and Why Users Stop Using a Digital Health Intervention to Promote an Active Lifestyle: Secondary Analysis With A Focus on Health Action Process Approach–Based Psychological Determinants (e30583)
Helene Schroé, Geert Crombez, Ilse De Bourdeaudhuij, Delfien Van Dyck. 246

User Control of Personal mHealth Data Using a Mobile Blockchain App: Design Science Perspective (e32104)
Arijit Sengupta, Hemang Subramanian. 261

Multipurpose Mobile Apps for Mental Health in Chinese App Stores: Content Analysis and Quality Evaluation (e34054)
Xiaoqian Wu, Lin Xu, PengFei Li, TingTing Tang, Cheng Huang. 281

Russian-Language Mobile Apps for Reducing Alcohol Use: Systematic Search and Evaluation (e31058)
Anna Bunova, Veronika Wiemker, Boris Gornyi, Carina Ferreira-Borges, Maria Neufeld. 294

XSL•FO
RenderX

Prioritization of Quality Principles for Health Apps Using the Kano Model: Survey Study (e26563)
Christin Malinka, Ute von Jan, Urs-Vito Albrecht. 305

Pulse Oximeter App Privacy Policies During COVID-19: Scoping Assessment (e30361)
Rachele Hendricks-Sturrup. 341

A Smartphone App to Improve Oral Anticoagulation Adherence in Patients With Atrial Fibrillation: Prospective Observational Study (e30807)
Keitaro Senoo, Tomonori Miki, Takashi Ohkura, Hibiki Iwakoshi, Tetsuro Nishimura, Hirokazu Shiraishi, Satoshi Teramukai, Satoaki Matoba. . .
3 5 0

Adherence to Growth Hormone Treatment Using a Connected Device in Latin America: Real-World Exploratory Descriptive Analysis Study (e32626)
Aria Assefi, Paula van Dommelen, Lilian Arnaud, Carlos Otero, Luis Fernandez-Luque, Ekaterina Koledova, Luis Calliari. 359

Attitudes Toward Mobile Apps for Pandemic Research Among Smartphone Users in Germany: National Survey (e31857)
Lorina Buhr, Silke Schicktanz, Eike Nordmeyer. 372

App-Based Relaxation Exercises for Patients With Chronic Neck Pain: Pragmatic Randomized Trial (e31482)
Daniel Pach, Susanne Blödt, Jiani Wang, Theresa Keller, Beatrice Bergmann, Alizé Rogge, Jürgen Barth, Katja Icke, Stephanie Roll, Claudia Witt. 389

Co-design of a Smartphone App for People Living With Dementia by Applying Agile, Iterative Co-design Principles: Development and Usability Study (e24483)
Sarah Fox, Laura Brown, Steven Antrobus, David Brough, Richard Drake, Francine Jury, Iracema Leroi, Adrian Parry-Jones, Matthew Machin.
4 0 4

Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study (e30557)
Aditya Vaidyam, John Halamka, John Torous. 416

Daily Level Association of Physical Activity and Performance on Ecological Momentary Cognitive Tests in Free-living Environments: A Mobile Health Observational Study (e33747)
Zvinka Zlatar, Laura Campbell, Bin Tang, Spenser Gabin, Anne Heaton, Michael Higgins, Joel Swendsen, David Moore, Raeanne Moore. . .
4 3 3

XSL•FO
RenderX

Review

Human-Centered Design of Mobile Health Apps for Older Adults: Systematic Review and Narrative Synthesis

Zethapong Nimmanterdwong[1], MD; Suchaya Boonviriya[2], BNS, MSc; Pisit Tangkijvanich[1], MD, PhD

[1]Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

[2]Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Corresponding Author:
Zethapong Nimmanterdwong, MD
Department of Biochemistry
Faculty of Medicine
Chulalongkorn University
Padtayapatana Building, 11th Floor
1873 Rama IV Road, Pathumwan
Bangkok, 10330
Thailand
Phone: 66 868622188
Email: zethapong@gmail.com

Abstract

Background: The world is aging. The number of older patients is on the rise, and along with it comes the burden of noncommunicable diseases, both clinical and economic. Attempts with mobile health (mHealth) have been made to remedy the situation with promising outcomes. Researchers have adopted human-centered design (HCD) in mHealth creation to ensure those promises become a reality.

Objective: This systematic review aims to explore existing literature on relevant primary research and case studies to (1) illustrate how HCD can be used to create mHealth solutions for older adults and (2) summarize the overall process with recommendations specific to the older population.

Methods: We conducted a systematic review to address the study objectives. IEEE Xplore, Medline via Ovid, PubMed, and Scopus were searched for HCD research of mHealth solutions for older adults. Two independent reviewers then included the papers if they (1) were written in English, (2) included participants equal to or older than 60 years old, (3) were primary research, and (4) reported about mHealth apps and their HCD developments from start to finish. The 2 reviewers continued to assess the included studies' qualities using the Mixed Methods Appraisal Tool (MMAT). A narrative synthesis was then carried out and completed.

Results: Eight studies passed the eligibility criteria: 5 were mixed methods studies and 3 were case studies. Some studies were about the same mHealth projects with a total of 5 mHealth apps. The included studies differed in HCD goals, target groups, and details of their HCD methodologies. The HCD process was explored through narrative synthesis in 4 steps according to the International Standardization Organization (ISO) standard 9241-210: (1) understand and specify the context of use, (2) specify the user requirements, (3) produce design solutions to meet these requirements, and (4) evaluate the designs against requirements. The overall process and recommendations unique to older adults are summarized logically with structural order and time order based on the Minto pyramid principle and ISO 9241-210.

Conclusions: Findings show that HCD can be used to create mHealth solutions for older adults with positive outcomes. This review has also summarized practical HCD steps and additional suggestions based on existing literature in the subfield. However, evidence-based results are still limited because most included studies lacked details about their sampling methods and did not set objective and quantifiable goals, leading to failure to draw significant conclusions. More studies of HCD application on mHealth for older adults with measurable design goals and rigorous research strategy are warranted.

(JMIR Mhealth Uhealth 2022;10(1):e29512) doi:10.2196/29512

XSL•FO
RenderX

KEYWORDS

connected health; development; elderly; HCD; mHealth; older adults; review; telehealth; UCD; usability; design; human-centered; app

Introduction

Background

The word "mHealth," or "mobile health," has been rising in popularity. A search of the term in an academic research database bears tens of thousands of results in 2020 alone. It is being studied as a medical intervention for arthritis [1], asthma [2], cancer [3], cardiovascular diseases [4], chronic kidney diseases [5], diabetes [6], multiple sclerosis [7], and various psychiatric diseases [8]. The idea of health care through mobile technology indeed accounts for its reputation. The World Health Organization defines mHealth as the "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [9]. mHealth is often brought up together with its broader term *telehealth* or *telemedicine*, which essentially means the practice of any kind of medicine with the help of technology across the distance [10]. With the COVID-19 pandemic, where social distancing is key, such digitalization of health care is becoming more relevant than ever [11].

mHealth and telehealth are the means to achieve timely and accurate health management; they help enable a seamless sharing of medical information between all those involved, creating the so-called connected health environment that the current trend strives for [12]. Successful integration of such innovations is believed to ensure universal health coverage, reduce health care costs, and improve clinical outcomes [9]. There were 5.2 billion mobile phone users at the end of 2019 with the estimation that the number will reach 5.8 billion by 2025, roughly 70% of the entire human population [13]. Diffusion of health care through a mobile medium in such a large populace will surely guarantee impact on a global scale. Real-world mHealth implementations across the globe are committed to educating patients, offering easier access to medical care, improving medical data storage and transfer, empowering health care providers, and boosting the efficiency of its institutions [14]. The synthesis of clinical evidence in the field is also on the rise. A meta-analysis of 11 lifestyle modification apps reported a significant reduction in the mean HbA_{1c} of the users in both short- and long-term observations [15]. Self-management interventions in 24 studies were shown to be able to decrease both systolic and diastolic blood pressures in patients with hypertension [16]. One systemic review that focused on pediatric asthma management reported increased treatment adherence in 13 studies, reduced exacerbations in 5, and improved quality of life in 4 [17].

Although mHealth has remarkable potential, most projects cannot scale to their own target population and fail to achieve the intended results. This can be attributed to (1) poor understanding of the end users and (2) failing business models [18]. Barriers to user adoption of mHealth can range from an individual level to a higher level of the policy governing its use. However, while policy barriers tend to impede new innovations or hinder the successful ones from a larger adoption [19], user-related barriers pose a more tangible challenge as no one might use the technology in the first place. A survey in the United States showed that about half of those who have downloaded health apps stop using them eventually [20]. The cause of this begins when inadequate user involvement makes it impossible to draft concise software requirements [21], which results in poor user acceptance and failure to scale [22].

These issues get even more complicated with older adults. The United Nations defines older persons as those aged over or equal to 60 or 65 years; now, over 703 million people are aged over 65 years, and that number is projected to double by 2050 [23]. Moreover, about 2 out of 3 older adults suffer from multiple chronic diseases [24], a condition to which mHealth proves to be a highly possible solution [3-7]. A myriad of frameworks and techniques have been employed to ensure the success of mHealth development and implementation with varying outcomes. Suggestions from research up to date stress the importance of having an in-charge multidisciplinary team working together with real end users rather than giving them the finished product out of the blue [25]. The International Organization for Standardization (ISO) 9241-210 further elaborates this concept in the term "human-centered design" (HCD) as the "approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques", in which the word "human-centered" is used to highlight that the process includes all stakeholders and not just the users [26]. Thus, in this review, the term "user centered" will be referred to as "human centered" to reflect its definition better.

Review Objective and Question

In searching for the best methodology to create the most usable mHealth, many have put the said value at the core of their work: having the humans at the center of focus. This review aims to explain how HCD can be applied to create mHealth suitable for older adults and to summarize the overall process with recommendations from relevant primary research studies of mHealth design and development.

The research question of this review is the following: How can HCD be used to create mHealth solutions for older adults? This issue was formed during the first author's attempt to develop an mHealth app for older adults to solve their current pain points in a geriatric wellness clinic. Despite the constant mentioning of HCD, previous scoping searches of literature bear a heterogeneous group of research studies differing in interpretation, execution, and the extent of evaluation. The need for further clarification on the procedural details is identified.

XSL•FO
RenderX

Methods

Design

A systematic approach following Siddaway et al's guide [27] was employed to ensure a robust acquisition of the existing literature related to the topic with a method as reproducible, transparent, and unbiased as possible. The review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [28] (Multimedia Appendix 1). Detailed methods are described in the review as no prior protocol was published.

Eligibility Criteria

Textbox 1 presents the eligibility criteria. As this review aims to draw from studies of a relatively new and emerging subfield of study, the criteria are inclusive. However, a certain degree of clarity in participants, qualitative or quantitative methods, analysis of the results, and discussions of the implementation results are required. Moreover, to best answer the review question, the included studies have to have these 3 key steps starting from (1) designing solutions based on existing problems, (2) developing the designed solutions, and (3) evaluating the developed solutions, all stated to be conducted in accordance with the HCD philosophy.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Community, primary, secondary, or tertiary care.
- Any qualitative, quantitative, or mixed methods study of original primary research.
- Participants must include, but not limited to, older adults (aged ≥60 years).
- Design goals must focus on mobile health (mHealth) solutions in the form of mobile apps intended for older adults.
- Study procedures must be in line with the human-centered design (HCD) philosophy.
- Studies must include details of mHealth apps and their development process, participants, design goals, and some implementation data.
- Studies depicting different processes of the same product/project are included. For example, an mHealth project might have 2 separate papers such as 1 for design and 1 for evaluation; both are included in this review.
- Trial and pilot studies are included.

Exclusion criteria

- Non-English language papers.
- Any type of literature review, narrative review, or systematic review.
- Studies with no relevant data or information that is of interest to the review question.

Search Strategy and Study Selection

Systematic searches were conducted from the following 4 databases: IEEE Xplore, Medline via Ovid, PubMed, and Scopus. To best ensure comprehensive search results, search strings were compiled from keywords of the review question. Listed below are those strings with their corresponding similar terms:

- "mHealth" OR "mobile health", for the app to be reviewed;
- "human centered" OR "human centered" OR "user centered" OR "user centered", the approach in question;
- "design" OR "development", the process required;
- "usability", an outcome of HCD according to ISO 9241-210;
- "elderly" OR "older adults" OR "geriatric", the target population.

Each group of strings was put together with the "AND" Boolean operator in the search engines as all of the above key terms were required by the set eligibility criteria. No date range was set. Manual searches on Google Scholar and the references of the eligible papers were also conducted to identify possible additional relevant papers for screening. All searches were performed by a single reviewer (ZN) on the same day (November 12, 2020). The reason why the ACM Digital Library was not included is discussed in the "Limitations" section.

Microsoft Excel was used to record and manage the search results; duplications were removed. Two independent reviewers (SB and ZN) screened the deduplicated results by titles and abstracts. The full-text screening was done by the same reviewers using the eligibility criteria from Textbox 1. The results were in agreement. The reviewers then proceeded to appraise the study qualities using the Mixed Methods Appraisal Tool (MMAT) for mixed methods studies [29]. Disagreements were resolved through discussions. As this review aimed to be inclusive, study quality was not used to exclude any paper from the review but rather to inform about the present research quality of the existing literature of interest. We chose MMAT as our appraisal tool because (1) it can appraise the heterogeneous methodologies of design studies and (2) its methodological focus helps reflect on the existing research critically. Table 1 presents the qualities of the included studies appraised by MMAT.

XSL•FO
RenderX

Table 1. Quality appraisal of included studies.

Studies	Criteria from the Mixed Methods Appraisal Tool														
	1.1[a]	1.2[b]	1.3[c]	1.4[d]	1.5[e]	4.1[f]	4.2[g]	4.3[h]	4.4[i]	4.5[j]	5.1[k]	5.2[l]	5.3[m]	5.4[n]	5.5[o]
Cornet et al [30]	1	1	1	1	1	N/A[p]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Cornet et al [31]	1	1	1	1	1	0	0	1	0	1	1	1	0	1	0
Fortuna et al [8]	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0
Harte et al [32]	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1
Harte et al [33]	1	1	1	1	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Petersen et al [34]	1	1	1	1	1	0	0	1	0	0	1	1	1	1	0
Srinivas et al [35]	1	1	1	1	1	0	0	0	0	1	0	1	1	1	0
Stara et al [36]	1	1	1	1	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

[a]Is the qualitative approach appropriate to answer the research question?
[b]Are the qualitative data collection methods adequate to address the research question?
[c]Are the findings adequately derived from the data?
[d]Is the interpretation of results sufficiently substantiated by data?
[e]Is there coherence between qualitative data sources, collection, analysis, and interpretation?
[f]Is the sampling strategy relevant to address the research question?
[g]Is the sample representative of the target population?
[h]Are the measurements appropriate?
[i]Are the confounders accounted for in the design and analysis?
[j]Is the statistical analysis appropriate to answer the research question?
[k]Is there an adequate rationale for using a mixed methods design to address the research question?
[l]Are the different components of the study effectively integrated to answer the research question?
[m]Are the outputs of the integration of qualitative and quantitative components adequately interpreted?
[n]Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?
[o]Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?
[p]N/A: not applicable.

Data Extraction

One independent reviewer (ZN) performed data extraction from the 8 eligible papers. The information from 5 mixed methods studies included (on the data extraction form) the year of the study, the country of the study, the name of the project (if stated), study design, design goals, participants, study methods, quantitative or qualitative data used, results, and key discussions. The information from the other 3 case studies included the year of the study, the country of the study, goals, and results. All extracted texts were manually typed in Microsoft Excel.

Synthesis of Results

Because of the heterogeneous nature of the included studies, narrative synthesis was chosen. Following Popay et al's guide [37], the narrative data synthesis was performed iteratively between the 4 key elements as explained in Textbox 2.

XSL•FO
RenderX

Textbox 2. Key elements for the narrative data synthesis.

- Developing a theory of how the intervention works, why, and for whom

Previous studies were carried out under the same hypothesis that human-centered design (HCD) helps make a more usable system for its users. This review adopted that same assumption and aimed to elaborate on how HCD works, especially for older adults, in steps.

- Developing a preliminary synthesis of findings of included studies

Textual descriptions together with tabulation were chosen to summarize and display the extracted data. A recurring concept was identified across the studies: the HCD process. To ensure transparency, suggested HCD activities from ISO 9241-210 were chosen to categorize these patterns into 4 steps as follows: (1) understand and specify the context of use, (2) specify the user requirements, (3) produce design solutions to meet these requirements, and (4) evaluate the designs against requirements [26].

- Exploring relationships in the data

Qualitative case descriptions were used to explore details and findings among included studies that correlate with each theme/step. A conceptual diagram was then created to answer the review question. The diagram was structured according to the Minto pyramid principle, using the following rules: (1) ideas at any level in the pyramid must always be summaries of the ideas grouped below them, (2) ideas in each grouping must always be the same kind of idea, and (3) ideas in each grouping must always be logically ordered [38].

- Assessing the robustness of the synthesis

All included studies were appraised by Mixed Methods Appraisal Tool (MMAT), and the synthesis process was reflected on critically.

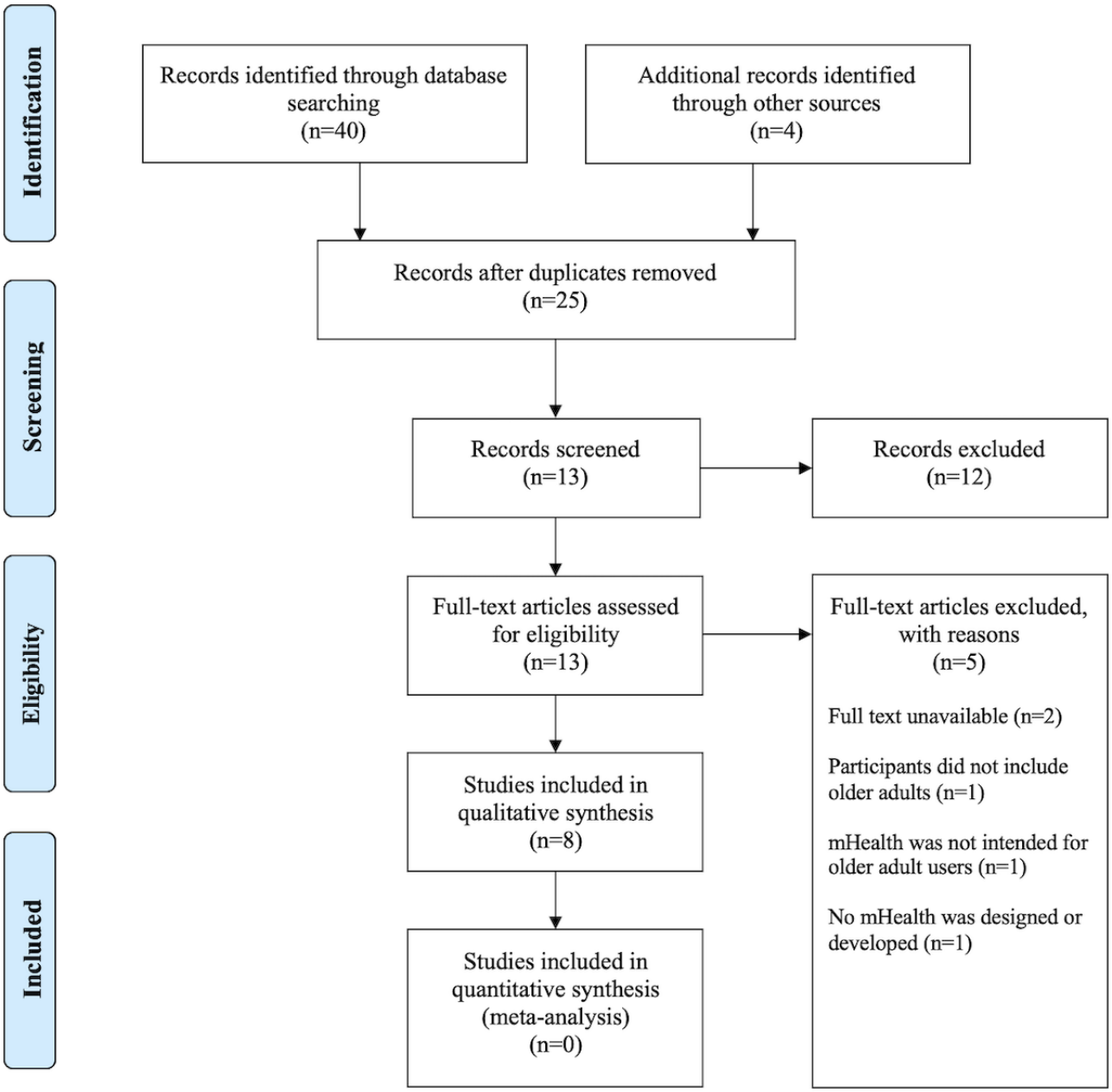
Results

Study Selection

Figure 1 shows the selection process of the included studies. The initial search yielded 44 studies, of which 40 were from the 4 databases and the other 4 were from Google Scholar. A total of 25 studies remained after the removal of duplications. Two independent reviewers (SB and ZN) screened titles and abstracts according to the criteria. The remaining 13 full-text studies were then assessed by the same 2 separate reviewers for eligibility. Five studies were excluded, as shown in Figure 1. Eventually, 8 studies were retained for this systematic review.

XSL•FO
RenderX

Figure 1. PRISMA flow diagram. mHealth: mobile health; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Study Characteristics

Five studies were mixed methods, employing both quantitative and qualitative techniques to design and evaluate mHealth apps. Three were qualitative case studies focusing on describing the methodology and problems found during the process with little or no actual quantitative or qualitative data from the research shown. One of the 3 case studies was about the same process from the same project of another included mixed methods study; it was included for its qualitative reflections on the design and development processes. All studies stated clear aims of the research and were deemed relevant to this review question.

To illustrate the overall quality of the included studies, their quality appraisal scores were reported in the MMAT-suggested format [39] in Table 1. Of the 8 included studies, all were rated to have adequate quality in their qualitative part. However, only 1 study had a passable rating of 60% in its quantitative part, while the rest were rated poor. The quantitative criteria that all studies failed were about the sampling strategy and the account for possible confounders. The description of the processes was lacking, leading to questionable results and interpretation. This issue was further explored in the narrative synthesis of results.

Table 2 summarizes HCD processes of the 5 included mixed methods studies. Four of the studies were conducted in the United States, with only 1 study conducted in Ireland. Two studies under the same project focused on patients with heart failure [31,35]. One study aimed at patients with psychiatric disorders [8]. The other 2 studies dealt with fall risk assessment and detection [32] and sarcopenia prevention [34], respectively. In addition, of the 3 included qualitative case studies, 2 reflected on the same project as the mixed methods study dealing with falls in Ireland [33,36]. By contrast, the remaining 1 study reflected on a different project targeted at patients with cardiac implantable electronic devices (CIEDs) in the United States [30].

XSL•FO
RenderX

Table 2. Summary of the included mixed methods studies.

Study	Project	Setting	Design goal	Participants	Methods
Cornet et al [31]	Engage	Academic health center, the United States.	To evaluate and test the usability of a self-managing heart failure system for older adults developed in a study by Srinivas et al [35].	(1) 13 older adults and (2) 2 caregivers	• Study I: (1) A structured interview was used to assess participants' daily self-management routines and technology familiarity. (2) The think-aloud method was employed as each participant completes 8 given tasks on the system. (3) Feedback from the patients after they finish was used. • Study II: (1) The system was re-designed after Study I. (2) A structured interview was used. (3) The think-aloud method was employed as each participant completed a given scenario in which he/she was to act as if he/she were the assigned fictitious character. (4) Feedback from the patients after they finish was used. (5) SUS[a] was used after usability evaluations. (6) NASA-TLX[b] was used after usability evaluations.
Fortuna et al [8]	__[c]	Specialized center, the United States.	To incorporate an existing psychosocial intervention into a selected mobile platform.	Phase I and Phase II: (1) authors; Phase III: (1) older adults and (2) experts; and Phase IV: (1) 10 middle-aged and older adults	• Phase I: (1) A literature review was done to identify requirements. • Phase II: (1) A literature review was done to find a suitable existing mobile platform. • Phase III: (1) The interdisciplinary panel of end users and experts work together to incorporate an existing psychosocial intervention into the chosen mobile platform. • Phase IV: (1) The think-aloud method was employed as each older adult goes through task-based usability testing. (2) Feedback from patients was collected. (3) Surveys based on SUS; Post-Study System Usability Questionnaires; and USE[d] questionnaires were used after each usability testing. (4) The ability to perform tasks without help was recorded in percentage.

XSL•FO
RenderX

Study	Project	Setting	Design goal	Participants	Methods
Harte et al [32]	Wireless Insole for Independent and Safe Elderly Living	Academic health center, Ireland.	To develop, assess, and enhance usability and user experience of a mobile app of a connected health system designed for fall risk assessment and fall detection.	Phase I: (1) 10 experts and (2) 12 older adults; Phase II: (1) 10 experts from Phase I; and Phase III: (1) 10 older adults from Phase I	• Phase I: (1) Likert scales were used to rate mock-ups called use cases. (2) The think-aloud method was employed during use case analysis. (3) Self-reported measures of the experts were collected. (4) Visual perception and cognitive processing metrics of older adults were collected. • Phase II: (1) Likert scales were used to rate the paper prototypes based on use cases by experts. (2) ASQ[e] and chosen usability metrics were used to rate the developed mobile working prototypes by experts after scenario-based usability testing. (3) The think-aloud method was employed during experts' mobile working prototype runs. • Phase III: (1) Likert scales were used to rate the mobile working prototypes by older adults. (2) ASQ, SUS, NASA-TLX, and chosen usability metrics were used to rate the mobile working prototypes by older adults after scenario-based usability testing. (3) The think-aloud method was employed during older adults' working prototype runs.
Petersen et al [34]	—	Academic health center, the United States.	To create a mobile app for older adults to monitor their use of a Bluetooth-connected resistance band for sarcopenia prevention.	Round 1: (1) 6 older adults; Round 2: (1) 3 clinicians and (2) 4 older adults; Round 3: (1) 3 clinicians and (2) 6 older adults	• Round 1: (1) Semistructured interviews gave information on how the app can be of use. • Round 2: (1) The think-aloud method was employed as participants go through the wireframes. (2) A verbal prompting method was employed to encourage participants to give their thoughts. (3) Oral feedback from participants was recorded as they go through the video contents to be used in the prototype app. (4) The SUS was used after each participant finishes. (5) The USE score was used after each participant finishes. • Round 3: (1) The think-aloud method was employed as participants go through the wireframes. (2) A verbal prompting method was employed to encourage participants to give their thoughts. (3) The SUS was used after each participant finishes. (4) The USE score was used after each participant finishes.

XSL•FO
RenderX

Study	Project	Setting	Design goal	Participants	Methods
Srinivas et al [35]	Engage	Specialized center, the United States.	To design, develop, and evaluate a consumer-facing health information technology system that supports heart failure self-care.	Phase I: (1) 63 older adults, (2) 35 caregivers, and (3) additionally data on 66 patients obtained from other literature; Phase II: (1) experts; Phase III: (1) 1 expert and (2) 5 older adults	• Phase I: Major themes of the app were synthesized from data gathered through direct observations at patient outpatient visits, standardized surveys on patient self-care, patients' electronic medical record reviews, and semistructured interviews focused on patient self-care. • Phase II: Core activities of the app were determined through educating, brainstorming, and design sessions of the research team. • Phase III continues in Cornet et al (2017) [31]: (1) heuristic evaluation done by the team's expert identified and classified usability flaws. (2) Structured interviews focusing on patients' self-care routines were done before usability testing. (3) The think-aloud method was employed during laboratory-based usability testing of the developed prototype as each older adult goes through the tasks given on a mobile. (4) Questionnaires adapted from the SUS were used after each usability testing.

[a]SUS: System Usability Scale.

[b]NASA-TLX: NASA-Task Load Index.

[c]Not stated.

[d]USE: Usefulness, Satisfaction, and Ease of Use.

[e]ASQ: After Scenario Questionnaire.

HCD Activities in mHealth Development for Older Adults

Overview

All 5 mHealth projects, from the included 5 mixed methods studies and 3 case studies, have the 4 key steps from ISO 9241-210 in their HCD processes, albeit described and mentioned to varying degrees. This section explores and illustrates these recurring steps across all included studies using the qualitative case description technique. All 8 studies are summarized and described in 4 HCD steps. Each step has 5 paragraphs representing a total of 5 mHealth projects: the first for patients with heart failure [31,35], the second for patients with psychiatric disorders [8], the third about falls in the elderly [32,33,36], the fourth for sarcopenia prevention [34], and the fifth about CIEDs [30].

Step 1: Understand and Specify the Context of Use

Understanding the context of use such as the end users, their current tasks, key activities, and working environment is essential to the design process; it helps guide how solutions should be tailored and set practical goals for the project [26].

Srinivas et al [35] used various HCD frameworks to develop an mHealth app that helped older patients with heart failure to improve their self-care engagement, health behaviors, and knowledge of the disease. In 2 years, the researchers collected data from 65 older patients with heart failure and 35 caregivers through direct observations at outpatient clinics, electronic medical record reviews, and semistructured interviews; the patients' health care routines, health literacy, environments, and supports were the priority. They conducted these field-based investigations in an academic medical center in Southeast United States. In addition, the authors included 66 other patient data from the United States and Singapore in an urban emergency setting. Details on the sampling method and rationales for the number of patients were not provided. It was also noted that not all data could be utilized fully in the design process. The qualitative quality of this study was adequate.

Fortuna et al [8] aimed to integrate self-management intervention into a mobile app for middle-aged and older patients with psychiatric disorders to promote self-care for better health outcomes. However, rather than obtaining data directly from potential users, the researchers gathered rationales and pain points of the project from a literature review. Details about the method were not specified in the paper. No quality appraisal of the included literature was presented. They then used the review results in the subsequent design. For example, integrating an existing intervention to an existing mobile platform was chosen over developing a new one because it was more practical. Characteristics specific to the elderly such as declining cognitive functions affecting their self-management and motivation, multimorbidity, and limited digital literacy were considered. The researchers also decided the intervention to be implemented based on the literature review: Integrated Illness Management and Recovery (I-IMR), an evidence-based medical practice for

XSL•FO
RenderX

psychiatric patients, was cited to be promising and thus chosen. The authors reported successful implementation and noted that identifying the unique needs of the intended users to guide the design process helped build a more usable product. The qualitative methodologies were appraised to be of adequate quality.

Stara et al [36] integrated HCD into the development process of their connected health system: the Wireless Insole for Independent and Safe Elderly Living (WIISEL), consisting of a pair of chargeable insoles with Bluetooth transmission, its charger, a smartphone app, a gait analysis desktop application, and an administrative web application. The authors drafted a preliminary concept of the system and then discussed it in 3 focus groups with 6 older adults and 6 stakeholders in each group; the focus groups were conducted at 3 separate sites: a primary care center in Ireland, a tertiary care center in Israel, and a specialized center in Italy. The sampling method was not specified. The qualitative quality is adequate. The authors concluded that barriers to technology-enabled care acceptance in older adults were related to security, intrusiveness into their home environment, lack of control, confidentiality, and usability issues worsened by aging. Thus, involving users early in the process proves vital in crafting a health care technology that matches actual older adult user needs, with elderly friendly user interfaces and safety being a priority.

Petersen et al [34] used HCD to develop an mHealth app featuring exercise videos to work with a Bluetooth-connected resistance band that together would help health care providers monitor older adults' exercise progress for sarcopenia prevention. A convenience sampling method was used. Six older adults were recruited from a primary care clinic at an academic health center in the United States. The researchers then conducted semistructured interviews to assess the patients' general views regarding mHealth, their current activities, and their opinions of the Bluetooth-connected band and sample exercise videos. They further explored the participants' opinions in using technology to help with their exercise therapy. The quality of the study was appraised to be adequate. The participants had positive responses to the idea. All had experience using smartphones. Notes from these interviews were then used as key information to guide further design processes.

Cornet et al [30] implemented HCD in developing an mHealth app that shared the information stored in CIEDs of patients with heart failure with the patient themselves. In 3 months, 24 older patients with heart failure, 12 of whom had CIEDs, were recruited from a major health system in the Midwestern United States for semistructured interviews to gain context about their health decision-making processes. No sampling method was stated. The interview utilized 2 notable approaches: (1) the critical incident technique, which involves asking the interviewee to recall a particular past event to gain insights through their actions and experience at the time; and (2) the think-aloud method, which lets the interviewee talk about what he/she was currently doing or would do in a given event. The researchers then analyzed and synthesized the gathered data into 2 outputs: (1) personas, a design technique that groups users based on their behaviors; and (2) use-case scenarios (or as-is scenarios), another design technique that depicts how users make decisions in hypothetical situations. These outputs were then cross-checked with 2 patient advisors, older adult patients who volunteered to help with design, and a group of 7 clinician experts from the same major health system. The patient advisor meeting was held early to gain additional inputs and feedback to help the team make more relatable personas and use-case scenarios. The clinician meeting was held later and focused more on the validity and feasibility of the subsequent processes. Besides, direct observations at the CIED clinic and meeting with 2 cardiologists were also done with the same objectives. The methods were appraised to be adequate in quality. As the paper is a case study, challenges and recommendations by the authors were reported. First, logistics issues including but not limited to compensation, conflicts of interests, older adult limitations, patient data, recruitment criteria, and stakeholder meetings need to be addressed or consulted with professionals to ensure efficiency and efficacy. Second, stakeholders should be involved early in the design process, and their roles should be identified clearly in how active they would be; for example, it might just be getting informed about the process, giving their opinions to the team, or having specific tasks given to them. The authors also added that more roles are not always better, and stakeholder involvement should be carefully balanced. Third, an adequate recruitment method should be employed to secure a representative group of potential users. Also, a selection of stakeholders who work well with the development team is key. Fourth, direct and timely communication between development team members and relevant stakeholders is recommended, although it might be difficult to achieve at times.

Step 2: Specify the User Requirements

The second step of HCD focuses more on synthesizing further outputs from the first step. The goal is to derive what the users need to do and their objectives based on the gathered context and then set a clear statement of user requirements for the solution designs [26]. User requirements lay down the groundwork for how the product should be created and which performance or criteria should be measured to evaluate the product. These requirements are often created along with other requirements of the product such as the requirements of the system stating that the system needs to be able to do a certain task because it will help users accomplish their goals.

Srinivas et al [35] reported a successful translation of major themes from the gathered data and created a set of requirements for the subsequent design. Thematic analysis was done to identify user needs; it was concluded that the patients lack adequate health information and communication regarding their conditions and disease progresses, they are disengaged from their self-care due to the added burden, and they are not equipped with practical knowledge nor tools for optimal self-management. The authors then held educating sessions with the design team, composed mainly of experts from technical and HCD backgrounds, on the phenomena of interest through various media and means from the collected data. Next, brainstorming sessions were held. The requirements were derived from the previously identified major themes from the research: the system needed to be viewable by the patients and potentially their health care providers, simple to use,

XSL•FO
RenderX

complementing well with their self-care routines, and customizable. Details of the performance goals to be evaluated were not stated.

Fortuna et al [8] derived requirements from their literature review of user interfaces for older adults, I-IMR contents, and the interdisciplinary panel of end users and experts recruited from the site where the project was intended to be launched. I-IMR is a clinical psychological intervention that requires both health care providers and patients to work together in 10 training modules/sessions covering 4 topics on psychoeducation, behavioral tailoring, relapse prevention training, and coping skills training over 8-10 months. Adapting this face-to-face intervention and its contents into a suitable mobile experience for middle-aged and older patients was key to the project. Details about the performance goals were not stated.

Stara et al [36] stated in their study that user requirements for the WIISEL-connected health system were defined by the design team together with 18 older adults who were potential users and 18 stakeholders who were geriatricians, neurologists, nurses, and physical therapists in 2 sessions of focus groups. Details of the participants and the final user requirements and performance goals were not provided.

Petersen et al [34] concluded pain points from their previous research regarding the exercise videos as follows: specific movements were hard to identify from low-contrast backgrounds, and instruction sounds were not heard clearly. Participants also stated that big and clear repetition numbers would help them better keep track, feedback and instructions would help them finish the exercises at home, and tablets were preferred as they have large screens. The authors did not show an explicit statement of user requirements or performance goals in the paper.

Cornet et al [30] did not provide details regarding the process of writing user requirements in their case study.

Step 3: Produce Design Solutions to Meet These Requirements

This HCD step focuses on designing how the users interact with the system based on the requirements from the previous step [26]. HCD strived for the best user experience. The process needs to be iterative and flexible to address user needs and requirements that are often hard to identify completely in 1 cycle. The outputs from this step are also used to explain and communicate the design concepts with stakeholders, simulate possible scenarios of its uses, and ultimately specify how the system is to be developed.

Srinivas et al [35] created design solutions from the requirements specified. The team members raised diverging ideas from those requirements. They then worked together to converge those ideas into 4 main potential design solutions: (1) a short-term intervention of 30 days to encourage user adoption of the system, (2) an avatar representing the results of different self-care routines to teach users about cause and effect in a more engaging way, (3) a function that allowed users to set and keep track of their goals to promote health behaviors, and (4) a tool that helped enhance clinical visit experience to improve communication and collaboration with their health care providers. Finally, the team decided to develop the mHealth solution based on the 30-day intervention idea. The system was to have 3 main modules to serve all user requirements previously set: (1) LOG, for users to log their health information; (2) HINT, a collection of short materials about heart failure disease and self-care; and (3) GOAL, gamified daily goals for better self-care behaviors. User–system interaction and user interfaces were then developed in subsequent design sessions composed of 3-6 members of the design team. Paper prototypes, Microsoft PowerPoint wireframes, and software prototypes were developed successfully. The authors noted that although clinical experts were consulted from time to time during this step, stakeholders, including end users, did not really participate in the design process; logistics issues and not knowing how to involve older adults were accountable for this approach. The authors also added that their waterfall approach to development, meaning the design process was linear and required time before evaluation, caused delays in solving design problems that could be prevented if a more agile approach was adopted.

Fortuna et al [8] created scenarios of uses, user–system interaction, and user interfaces based on the identified user needs and requirements. The system was designed to have the following: an ability to be customized, a tracking system to show users their progress, a monitoring system to send data back to health care providers, and a messaging system from health care providers for more human interaction and a smoother workflow. The text contents from I-IMR manual, originally intended for clinicians, were also modified to fit smartphone pages and rewritten using the Flesch-Kincaid Grade Level formula in Microsoft Word to a simpler sixth-grade level. The 4 core topics of I-IMR were also re-designed to fit the mobile experience. Psychoeducation used short videos that showed clinicians teaching self-management techniques to patients to help users master the skills. Behavioral tailoring utilized educating modules, a medication schedule function, and a reminder system to make patients take their medication on time. Relapse prevention training, usually done by exploring a patient's experience and identifying triggers to create a prevention plan for a possible relapse, offered an already made plan that was accessible and editable on mobile at any time. Coping skills training also used videos as media to equip users with the tools to help them in the real world, for example, relaxation videos that guide users to self-soothe and calm themselves down. In addition, issues regarding data security and mHealth user disengagement were addressed: the project adhered to Health Insurance Portability and Accountability Act (HIPAA) compliance and involved health care providers to encourage adoption.

Stara et al [36] stated in their study that the subsequent design process of the smartphone app of the WIISEL system is elaborated in 2 studies included in this review: 1 mixed methods study focused on the findings by Harte et al [32] and the other case study focused on the HCD methodology also used by Harte et al [33,36]. Rapid development was employed to create, test, and produce 4 versions of prototypes, 2 on paper and 2 on mobile. The first paper prototypes (also called "use cases") consisting of scenarios of use, descriptive end user profiles, storyboards, and interface mock-ups were created from the

XSL•FO
RenderX

opinions of all project stakeholders. These use cases were based on key activities that the users needed to carry out. Ten multidisciplinary experts and 12 older adults were then recruited by a purposive sampling method to analyze these use cases. This analysis quickly identified usability problems that were in turn fixed by the development team. The first mobile prototypes together with user manuals and the updated, second paper prototypes were then produced. The manuals were created to help address usability problems that could not be fixed by the development team such as the built-in buttons, the operating system keyboard design, the impracticality of an automatic data sync, and the connection limitations. The same experts then evaluated the first mobile prototypes simultaneously with the guide of the second paper prototypes and the user manuals. The results from this second expert analysis were then used to design the second mobile prototype for another usability test with end users. The authors added that multiple inputs from the relevant stakeholders, although divergent in nature, are essential to HCD and could be obtained only with enough rounds of iterations. Thus, the rapid cycles of using paper prototypes and expert evaluation for fast feedback before end user usability testing are recommended.

Petersen et al [34] updated the exercise videos from user pain points and created mobile prototypes as black and white wireframes showing simple outlines of the designed user interfaces. The researchers then presented exercise videos and the wireframes with different design approaches to 3 clinicians and 4 older patients from an academic health center in the United States for additional inputs. The pain points gathered from the older patients about the videos and the wireframes were as follows: the video instructions should be slower and have more details, the videos should have subtitles with large fonts, the video sound frequency should be adjustable, and the progress bars in the wireframes should be vertical. By contrast, the clinicians were content and also suggested the use of the Borg Scale of Perceived Exertion to measure each exercise difficulty in comparison to the others. Finally, the team created the interactive mobile prototypes featuring playable updated exercise videos with clear instructions and the colored user interfaces for usability testing in the next step. The authors noted that health management is a process that needs both health care providers and patients; therefore, the more stakeholder groups involved, the more complete the design of the mHealth that aims to assist the process.

Cornet et al [30] produced and improved 4 versions of prototypes through iterative prototyping and testing. The mHealth app for patients with heart failure with CIEDs was designed to have 4 main features: a heart health score derived from CIED data for the patient audience, self-assessments covering topics of recommended self-care routines, guides for better heart failure self-management, and logs showing data from CIEDs. The first prototype design was reported to take 5 months to complete. The 2 patient advisors from the research phase helped review this early prototype design. The later 3 prototypes then took 2, 1, and 3 months, respectively, with the final prototype going through refinement for heuristic evaluation for another 1 week. Feedback from each prototyping and testing was used to improve the later versions. More details of the process were not stated in this case study; however, challenges and recommendations were reported. First, the authors found that design solutions should be based on evidence gathered from the potential users of the project to avoid bias or assumptions of the design team. This challenge benefits from stakeholder involvement and rapid cycles of testing and feedback. Second, design solutions and features should be prioritized and focused on. Grouping these solutions into modules and structurally planning how to develop and test them to determine what works and what does not could help simplify the process. The third point elaborates on the first point but focuses on the feasibility of the proposed design solutions: to balance the design team's creativity with practicality. All limitations or regulations regarding the mHealth and its implementation should be worked out properly with the stakeholders to avoid project failure.

Step 4: Evaluate the Designs Against Requirements

The human-centered evaluation activity is vital to HCD and is iterative by its nature [26]. As illustrated in the third step, producing design solutions usually follows by evaluating them to assess their abilities to fulfill the requirements, obtain user feedback, gain more user needs, and quantify the results as baselines or for comparisons.

Srinivas et al [35] conducted a series of evaluations on the wireframes and the prototypes in parallel with the design process. First, heuristic evaluation by the team's human–computer interaction expert guided by Nielsen's usability heuristics [40,41] was done before the software prototype development. This helped transition the static wireframes into the interactive software prototypes and identify usability flaws early in the process for correction. The authors reported 45 flaws, of which 6 were major flaws. The corrected software prototypes were then evaluated by older adults and caregivers as elaborated further in another study by Cornet et al [31]. The researchers conducted the evaluation in 2 phases in a laboratory setting: (1) a task-based usability test with 5 users and (2) a scenario-based usability test with 10 users. A total of 13 patients with heart failure and 2 informal caregivers aged over or equal to 60 years were recruited from an urban and another suburban outpatient cardiology clinic of an academic health system in the Midwestern United States. All consented to the study and were compensated with US $40 gift cards. Details of the sampling method were not specified. All participants were given mobile devices with the software prototypes installed ready for testing. Both tests involved structured interviews about users' self-care routines and familiarity with technology at the start, the think-aloud method by talking out loud about what they were thinking during testing, and the use of standardized evaluation tools at the end. The tools include (1) System Usability Scale (SUS) consisting of 10 questions about the overall usability of the product and reporting in a score of 0-100 with 68 defined as average usability [42] and (2) NASA-Task Load Index (NASA-TLX) consisting of 6 scales to assess the cognitive load expended during product use [43]. SUS was used in both phases, but NASA-TLX was used only in the second phase. The tests were video recorded. The software prototypes were updated between the 2 usability tests. It was reported that SUS rating improved in the second phase from below average to above average. However, the

XSL•FO
RenderX

authors stated that the result could be affected by the design changes made, the different usability testing methods, and the sampling techniques. Moreover, the wording of SUS was shown to be difficult to understand to a certain group of older adults and might not reflect the real usability of the system [44]. The methodologies were appraised to be of inadequate quality. The authors added that quantitative results from the standardized tests did not capture the whole picture of the system usability issues and should be interpreted together with the qualitative results. Some older adults also showed resistance toward these usability techniques, that is, the think-aloud method was strange and the fictitious event of the scenario-based testing was counterintuitive as they had to remember the mock details that were irrelevant to them and got distracted. Logistics issues such as the locations of the testing sites, the set ups of recording tools, and the transportation of the older participants also need to be addressed.

Fortuna et al [8] conducted 2 cycles of task-based usability testing with 2 different groups of 5 participants each. The authors deemed a minimum of 5 participants could identify most usability issues [45]. All participants were middle-aged and older patients with both medical and psychiatric illnesses recruited from 2 mental outpatient programs in New Hampshire. A purposive sampling method of reviewing medical charts and reaching out to potential patients for informed consent was used. Gift cards worth US $20 were provided upon participation. The participants were given mobile devices with the app installed and a list of tasks to complete. They were orientated on how to use the devices and what the think-aloud method was. The researchers also asked the participants about the user interfaces and assigned them adapted surveys based on SUS, Post-Study System Usability Questionnaire (PSSUQ), and the Usefulness, Satisfaction, and Ease of use (USE) questionnaire. PSSUQ is an 18-item questionnaire with 7 rating scales and 1 not-applicable rating, assessing user satisfaction with the system [46]. The USE questionnaire also contains multiple items with 7 rating scales that explore 3 dimensions: usefulness, satisfaction, and ease of use [47]. All sessions took approximately an hour and were audio recorded if allowed or noted in detail if not. Updates on the app were made between the 2 cycles from user feedback: the text and video contents were shortened and the reading level was reduced from sixth grade to fourth grade. The authors reported that all participants could finish given tasks independently and both the qualitative comments and quantitative surveys had positive results, suggesting the users were satisfied with the app and would continue to use it if encouraged to do so. The mixed methods methodology was appraised to be inadequate in quality. The authors remarked that (1) future behaviors, that is, whether the patients would use the app in a real-world environment or not, were hard to predict from 1 hour of usability testing in a controlled environment, (2) the patients recruited specifically for the purpose of usability testing might lack heterogeneity and did not fully represent the intended vulnerable group of interest, and (3) technology constraints of utilizing an existing platform were reported. The authors also added the results might prove relevant and beneficial to the research of a similar fashion, and more studies on the mHealth intervention effectiveness were needed.

Stara et al [36] incorporated evaluation early in their HCD process as shown in 2 studies by Harte et al [32,33]; the authors then conducted user testing of the finished system with 54 older users [36]. During the course of producing design solutions, evaluation was done on (1) the first paper prototypes or the use cases, (2) the second paper prototypes, (3) the first mobile prototypes, and (4) the second mobile prototypes. The participants were 10 multidisciplinary experts and 12 older adults recruited using a purposive sampling method. Self-reported measures regarding the experts' knowledge together with the older adults' visual perception and cognitive processing metrics were reported. For the first paper prototypes, 10 experts and 12 older adults analyzed the prototypes by going through each use case. The think-aloud method was used to gather qualitative inputs, and Likert scales that asked the users to rate the user interfaces and task flows were used after each use case to obtain quantitative results. Likert scales are 5-point scale questionnaires that can be used to quantify user satisfaction; the question can be, for example, "I have no problems using the system." Usability problems of the first paper prototypes were then identified from think-aloud transcripts and grouped according to a derived set of heuristics [48]; the problems were then given severity rating based on the results of the related use case Likert scores. The prototypes were then updated accordingly. For the second paper prototypes, the experts analyzed the updated use cases in the same manner again to compare them with the first ones: most usability problems were reported to improve. For the first mobile prototypes, scenario-based usability testing was done by the experts as if they had been first-time users. The experts were able to use user manuals during testing. The sessions were also video recorded. Think-aloud scripts together with After Scenario Questionnaire (ASQ), SUS, and 3 usability metrics (ie, time taken to complete task, errors made, and completion rate) were used to update the user manuals and the user interfaces. The ASQ is a 3-item questionnaire regarding ease of completion, time taken to complete, and support information of the system with a 7-point scale, where a lower score indicates greater satisfaction [49]. For the second mobile prototypes, both task- and scenario-based usability testing were done by 10 older adults. They had access to the user manuals during testing. The sessions were video recorded. Data were obtained using the same methods as the first mobile prototypes with the addition of posttest interviews about general impressions of the system and NASA-TLX. The authors reported that the system achieved acceptability among end users. This is the only included mHealth project that has adequate quality appraised by MMAT. The authors concluded in their study that (1) older adults needed clear feedback from the app due to technology unfamiliarity, but imposing feedback such as alerts or cautions should be used only when necessary to avoid anxiety for the same reason; (2) older adults were found to be uncomfortable with touchscreen keyboards, thus minimizing or simplifying them would be ideal; (3) standardized tests such as SUS, ASQ, and NASA-TLX proved to give concordant and valuable information regarding the system usability, but they should be interpreted together with the more objective metrics, such as time taken to complete task, errors made, and completion rate for more tangible results; (4) expert evaluation before end user usability testing was efficient and

XSL•FO
RenderX

thus recommended; and (5) multiple inputs from different stakeholders, despite being divergent in nature, were essential to HCD. The process took 12 months to complete, of which the first prototypes took around 6 months; it was noted that interviewing and testing all the participants in the first phase were the causes of the long duration. Finally, 54 older users then tested the WIISEL system, both the mobile app and the soles, in Ireland, Israel, and Italy [36]. The usability testing had 2 stages: (1) the 3-day pilot stage had 15 participants test the system in a laboratory setting for a day and then at home for 2 days without specific instructions, and (2) the validation stage had 39 participants use the system at home for 14 days also without specific instructions. The participants completed the 12-item Quebec User Evaluation of Satisfaction with Assistive Technology questionnaire (QUEST) and SUS after each stage. Both had positive results. The authors did not detail the process as the paper was a case study. They concluded that technology acceptance was most affected by the system effectiveness but could also be positively influenced by proper user training and support.

Petersen et al [34] evaluated the wireframes and the prototype app using both qualitative and quantitative methods. Three clinicians from an academic health center in the United States participated; 6 older patients were recruited from a primary clinic of the same health center. A convenience sampling method was used. Think-aloud and verbal prompting methods were employed during testing to gather qualitative feedback from the participants. SUS and USE questionnaires were used after each participant finished testing the wireframes and the prototype app. The USE questionnaires comprise 30 questions asking about usefulness, ease of use, ease of learning, and satisfaction of the system with a 7-point scale to rate them. All sessions were audio recorded. Usability scores of the wireframes and the prototype app were calculated together and showed no statistically significant differences between the clinician and the patient participant groups with mean SUS scores of 65.8 and 66.8, respectively. The mixed methods study was appraised to be of inadequate quality. In addition, sentiment analysis of the participants' recorded statements was done; its results were in accordance with the SUS scores. A further application of natural language processing–based Dirichlet allocation topic modeling of the recorded statements showed that clinicians and older patients had different topics of interest regarding the mHealth system. The authors concluded that (1) inclusion of different stakeholder groups was vital to HCD because each has a different perspective on the mHealth system as illustrated in the study, (2) sentiment analysis could prove useful to HCD by effectively and efficiently analyzing qualitative inputs alongside traditional usability techniques, and (3) future research on incorporating sentiment analysis and natural language processing in HCD was encouraged.

Cornet et al [30] conducted 3 usability evaluations and 1 heuristic evaluation concurrently with the iterative production of 4 prototypes. The 3 rounds of usability evaluations had 4, 8, and 12 older patients with heart failure as participants, respectively. Details of the sampling method were not specified. The software prototypes were installed on prepared smartphones for the tests. The first 2 rounds were 90-minute task-based usability testing. Participants had to complete demographic surveys and the Newest Vital Sign (NVS) health literacy screening before the test. NVS is a screening tool that takes 3 minutes to complete; it has 6 questions asking about a nutrition label that is given to the patient to assess their health literacy [50]. The think-aloud method was used during the test. Participants had to complete SUS, NASA-TLX, and user acceptance survey once finished. They were also interviewed about the system after the test. The third round was a 90-minute scenario-based usability testing simulating the use of the system in the first 10 days. The participants in this round were also required to have CIEDs. Data were gathered in the same manner as in the first 2 rounds. Finally, heuristic evaluation was done by 3 outside HCD experts for refinement of the system; the process took 2 weeks. Details of the results were not stated. However, the authors reported challenges and recommendations found as the results of their case study. First, laboratory usability testing is good for detecting general software issues (eg, user interfaces and navigation), but it might not be able to address real-world usability issues. Thus, system evaluation at the actual site of the intended setting should be considered as time and the budget allow. Second, standardized tests should be adapted to fit real users, for example, the word "cumbersome" in SUS was changed to "awkward" in the study as the older adults could understand it better. Third, the testing process tends to get complicated and lengthy with numerous tools and techniques employed, therefore the HCD team should opt to cut reducible workload, manage time between testing and analyzing, and look for the possibility of utilizing automated data collection or analysis.

Discussion

Principal Findings

This systematic review has shown how HCD can be used to create mHealth for older adults, with additional recommendations reported. Eight studies are included in this review: 5 are mixed methods studies and 3 are case studies. All studies were published recently starting from 2017 onward, suggesting that the subfield is relatively new. All were conducted in developed countries and mostly in academic or specialized health care settings. Because of the diverse methodologies and details of the included studies, we used the Minto pyramid principle and the 4 HCD steps from ISO 9241-210 to guide the creation of 3 conceptual models: Figure 2 shows a structure of HCD team members and stakeholders in the HCD process, and Figure 3 shows how HCD can be applied to create mHealth solutions for older users. The following discussion explains the models further and also explores limitations with recommendations for future research.

XSL•FO
RenderX

Figure 2. Pyramid model of HCD project by structural order. HCD: human-centered design.

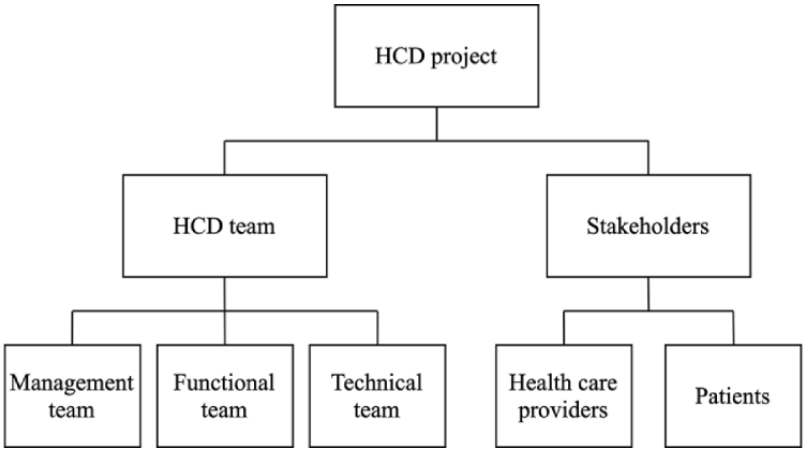
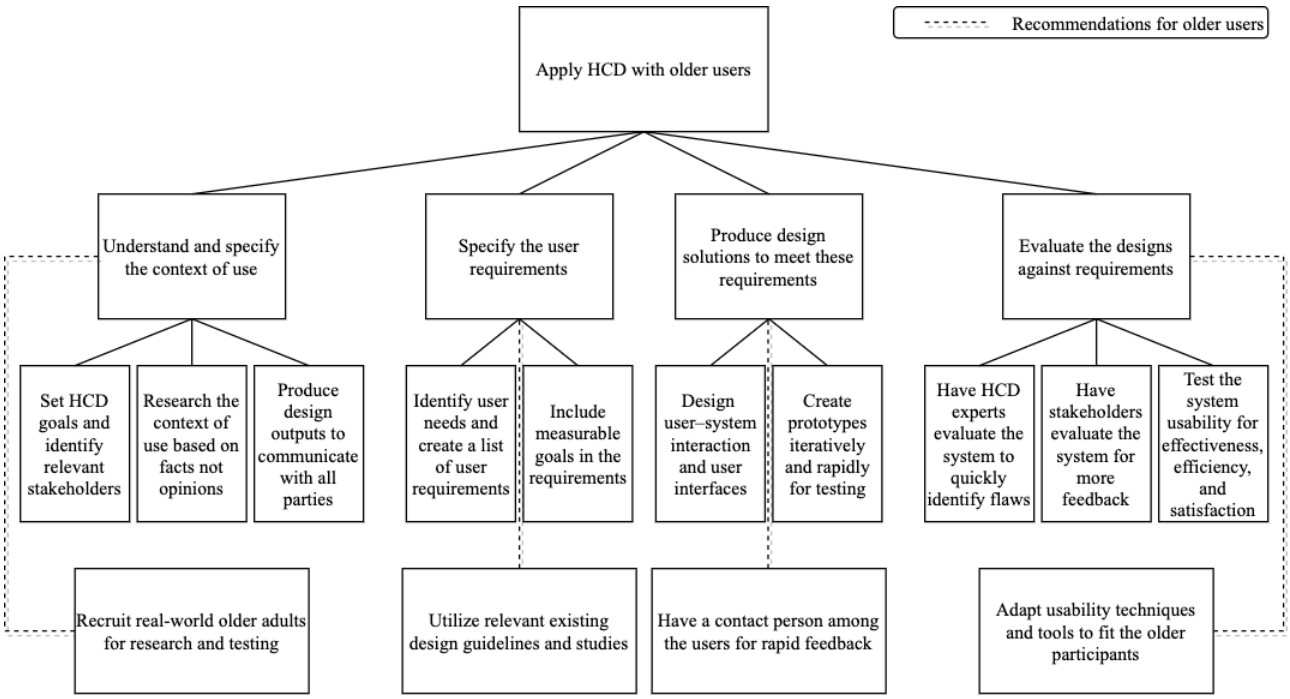


Figure 3. Pyramid model of HCD process with recommendations for older users by time order. HCD: human-centered design.



First, mHealth ideas, either novel or of existing concepts, should be based on what the users need, not what the creators want. As illustrated in the included studies, the authors, usually acting as the management team that oversees the project, identify and base their proposed mHealth solutions on real stakeholders both directly [30-36] and indirectly [8]. The first step of HCD investigates whether the solution fits well with the target users; this step also aims to produce outputs that ensure all HCD team members and stakeholders share the same vision. Stara et al [36] conducted focus groups to discuss their preliminary concept with relevant stakeholders. Researchers, or the functional team members, then recruit relevant stakeholders, both health care providers and patients, to gain more insight into their context of use: the users, their environment, and their current activities.

The included studies' details and rationales for the number of participants and the sampling method were diverse and vague. For example, Fortuna et al [8] had no participant and relied solely on a literature review to identify older adults' needs, whereas Srinivas et al [35] had a total of 100 participants and remarked that the gathered data proved to be more than they used for design. However, it should be noted that the 2 studies differ in their design goals, where the first wanted to implement a known intervention on mobile phones (app), while the second sought to identify new problems from an existing routine entirely.

Besides quantity, most included studies recommended that the sampling method include diverse groups of participants to ensure HCD solutions reflect real-world problems. Reaching out to older adults who are more physically inept or socially disadvantaged can be challenging. Fortuna et al [8] tackled this by building on the results of existing research in middle-aged and older adults with serious mental illness. Cornet et al [30] suggested recruiting key stakeholders who know how to approach such a group of patients to help. Harte et al [32] used purposive sampling and evaluated their participants' visual perception and cognitive processing to ensure the process was inclusive. Thus, the recruitment of older adults for HCD projects should be flexible and inclusive to best serve HCD goals.

XSL•FO
RenderX

Information on the context of use was mostly gathered through qualitative methods in the included studies. Observation is valued more than opinions in HCD: it shows how users currently pursue their goals from an unbiased perspective. Interviewing techniques that can be employed are (1) the critical incident technique, and (2) the think-aloud method of a fictitious scenario [30]. This factual information of the context of use is crucial to HCD as the functional team needs it to create HCD outputs such as personas and use-case scenarios to communicate with the technical team to develop a suitable mHealth solution. All 3 actions, which are (1) the management team setting design goals and identifying stakeholders, (2) the functional team gaining insight into the users, and (3) the functional team creating design outputs, can and should be done iteratively to truly understand the context of use.

Second, mHealth solutions need to address the current pain points of the users and ensure they achieve their intended goals; a clear understanding of user needs and a concise list of user requirements help the HCD team accomplish that. The context of use plays a vital role in identifying the user problems from their current activities and what the users need to solve them. Then, user requirements based on these user needs are created to guide the HCD team on how the solutions should be designed. These requirements can be obtained through a literature review [8] or a direct contextual inquiry of recruited participants [30,35]. For example, Srinivas et al [35] successfully derived user requirements from the user needs identified through thematic analysis of the established context of use.

The included studies also pointed out that there was a set of requirements unique to older users; however, most had not listed these requirements at the beginning and dealt with them only after the users raised the problems in usability testing. These design considerations for older adults are well-established: Harte et al [51] reported an extensive list of HCD considerations for connected health devices for older adults, and Li et al [52] identified barriers to mHealth adoption by older adults in their narrative review. With these guidelines, mHealth solutions can be designed to suit older adults' physical and cognitive limitations prior to testing for efficiency. Psychosocial factors such as motivation, technology perception, and social influence need to be addressed as well to ensure adoption.

Specifying user requirements also means setting measurable goals for the mHealth system. This usually requires gathering quantitative data for the context of use, such as the duration to complete the conventional I-IMR, which is approximately 8-10 months [8]; the HCD team might set the goals for their system to take only 4-5 months accordingly. If no goal is set, the way to assess the system in the subsequent steps will be limited.

Third, HCD seeks to create an ideal system through iterative prototyping together with the stakeholders, to make certain all user needs and requirements are accounted for. In the beginning, the functional team should design how the users will interact with the system and how the interfaces will be like. Harte et al [33] demonstrated this in their study by creating use cases as outputs to be analyzed by stakeholders for feedback. These outputs are called low-fidelity prototypes because they are easy to create, simple to change, and able to quickly convey the design concepts to all relevant parties. Besides, user interfaces can be based on existing design guidelines, such as the literature review about the unique user needs and requirements of older adults done by Fortuna et al [8]. Once the design is refined and approved, the technical team could then create an interactive software of the system or high-fidelity prototypes.

It should be noted that high-fidelity prototypes are not open for major changes or costly to do so; the best approach would be to finalize user–system interaction and user interfaces before their creation. As this HCD step requires iteration, an agile approach is recommended [35]. Agile is a software methodology based on rapid and iterative prototyping to gain continuous feedback from users, allowing developers to quickly create, evaluate, and improve their solutions to best fit the users [53]. Communication with stakeholders is key in this step, especially with the older population whose participation tends to be low due to their technology ineptness, physical and cognitive limitations, and logistical issues, for example, timing and travel [30]. Cornet et al [30] recruited patient advisors, older adult volunteers, to bridge these gaps: these patient advisors were able to give rapid feedback from users' perspectives. Due to HCD being agile in nature, this step is often done together with the evaluation step.

Fourth, evaluation comes after the production of design solutions. The functional team should work closely with the technical team to evaluate the produced solutions. In HCD, evaluating the designs is done by usability experts and end users. A number of included studies recommended system evaluation by usability experts before end users. Expert evaluation can identify and classify usability problems early in the process where changes are less punishing; it is also much simpler to arrange compared with its user-based counterpart [32,33,35]. The process can be conducted according to standards such as usability heuristics [30,32,33,35] or by having the experts role-play as real end users [32,33]. However, it should be noted that the greater the difference in usability knowledge between the experts and the users, the more divergent the results from the 2 groups might be [26].

Evaluation by end users is critical to the process because getting feedback from the target users and improving accordingly would surely make the system usable for them. This result can be obtained by recruiting the right participants. Fortuna et al [8] reported that their sampling method with the intention to select only willing participants to test the system could have excluded the group of real end users who might be less eager to participate. Harte et al [32] dealt with this problem by verifying that the recruited participants fit well with HCD goals: the participants' visual perception and cognitive processing were measured to confirm that the sampling method was inclusive enough. As for the number of participants, no standard has been agreed upon, but 5-10 participants are typically enough to discover major usability problems [45].

User-based evaluation methods range from giving the users specific tasks in a controlled environment to letting them use the system in the real world; the complexity also increased respectively so. Cornet et al [30] remarked in their study that usability testing in a controlled laboratory setting is more

XSL•FO
RenderX

prevalent in research as it is less complicated to set up; however, it also has limitations as the set time and place cannot replace the real intended context of use. Srinivas et al [35] discussed further in their study about the 2 laboratory-based methods: the task-based test is good for the identification of user interface flaws due to its straightforwardness in giving users a set of smaller tasks to complete, while the scenario-based test can help explore how the users perceive the system and its purpose in a similar way of using it in the real world. To summarize, laboratory-based usability testing is recommended during iterative prototyping, and researchers should then plan for usability testing in the real context of use if possible [30].

The included studies concurred that both qualitative and quantitative data should be interpreted together for robust usability evaluation results [8,30-36]. Qualitative data are gathered from interviews and participants' statements during the process. These statements can be encouraged by utilizing usability techniques such as the think-aloud and verbal prompting methods. Quantitative data are collected through standardized usability tools and usability metrics. However, the older population might show resistance toward these usability techniques and tools [35]. Orientation sessions about the goals of these techniques prior to testing could ease the older users' doubts [30], and adaptations of standardized tools such as using simpler synonyms and combining multiple tools into a single questionnaire could also help [8,30]. In addition to traditional means, Petersen et al [34] showcased the use of sentiment analysis and natural language processing to help analyze qualitative data in HCD; using such technology could improve the overall process, and more research is suggested.

ISO 9241-210 defines the components of usability as effectiveness, efficiency, and satisfaction [26]. Evidence on how HCD improves the usability of mHealth for older adults is still lacking as most included studies only report satisfaction based on SUS, with only 2 studies by Fortuna et al [8] and Harte et al [32] reporting objective usability metrics that represent effectiveness and efficiency of the system [8]. Setting baseline goals during the second step of specifying user requirements might help researchers draw more substantial conclusions.

Although limited, the positive outcomes from the studies in this review show that HCD can create usable mHealth systems for older adults. Stara et al [54] further suggested that this point held true even when the system was used in other cultural settings adjacent to the one it was developed in: the WIISEL system, which was developed in Ireland, also had good usability scores when tested in Israel and Italy. They added that these results meant the system's usability demands were within the capabilities of the users in the 3 countries. Human capabilities can be divided into 4 categories: physical, sensory, emotional, and intellectual [55]. The older population shares limitations in all these aspects, and by carefully addressing their needs with HCD, designers can create universally accepted products for older users across the globe [51].

The author again emphasize that user involvement in HCD is paramount to obtain such outcomes. Older adults are not extra design challenges to solve. Empathy toward users as individuals with pain points is essential to HCD; stereotypes and bias against older people could lead to design failures if left unchecked [56]. To avoid such pitfalls, we have to learn from the *untold stories* [30]. This review has gathered and summarized practical HCD challenges and strategies from primary research to aid HCD implementation with older adults.

Hastened by the COVID-19 pandemic, the field of mHealth will only expand. Moving forward, digital health solutions are aiming further than empowering patients and enhancing delivery. They are going for "digital therapeutics." These evidence-based interventions aim to prevent and manage medical conditions through digital platforms and mobile devices; one of its focuses is to deliver lifestyle therapy to combat chronic diseases such as type 2 diabetes [57]. Older adults are major target users as most have chronic conditions and can benefit greatly from these digital lifestyle therapies. However, the field is in need of solutions for effective development, testing, and deployment [58]. Future research on implementing HCD in digital therapeutics might be able to solve these issues and improve the health of the older population as a whole.

Limitations

Limitations of this systematic review are acknowledged. First, the ACM Digital Library was not included in this review despite being in the relevant field. We did search the database on the same day as the others: no studies from the ACM Digital Library passed our criteria. We then failed to mention this once we proceeded with the review. However, we ran another search with the same strategy on the database in May 2021 to recheck; 14 studies found did not pass our abstract screening according to our established eligibility criteria.

Second, the research question aims to address the whole HCD process, but an existing body of literature proves to be limited as the topic is an emergent subfield, especially with older adults as the target group. Although the criteria are forgiving, the search strategy and the inclusion criteria still demand that all steps of HCD are implemented in each app development. This excludes a large number of studies that feature only a part of HCD. For example, one study might focus on qualitative interviews without applying them, while another might test a newly developed system that is based solely on the authors' vision, not actual user needs. Nevertheless, the included studies complement one another and thus can accommodate the research question as illustrated in this review.

Third, the highly diverse HCD goals and methodologies in the included mHealth apps restrict the means of analysis and synthesis of results. All studies relied heavily on various qualitative means for HCD such as literature reviews, interviews, and field notes from direct observations. Even the seemingly same approaches, such as interviews, still differ in detail such as the time, the duration, the focus, and the questions. Most studies also focus more on the process not the result, or in the case of case studies, the process itself is the result. This might be due to the fact that baselines of the existing activities are not established in the second step of HCD, specifying the user requirements, so comparisons for effectivity and efficiency of the newly developed mHealth interventions cannot be made with objective metrics. Because of that fact, the included studies have to be reviewed with qualitative techniques using narrative

XSL•FO
RenderX

synthesis and guided by ISO 9241-210 together with logical ordering of the Minto pyramid principle [26,37,38]. Quantitative results of the included studies, which are based on the less tangible satisfaction results of standardized tools and often lack a definite conclusion, are also underutilized.

This leads to the fourth limitation regarding the included studies: all but 1 of the 5 mixed methods studies are rated to be of inadequate quality by the MMAT. Their quantitative components lack clarity. They do not explain their sampling methods or have done so insufficiently, resulting in the inability to deem their samples representative of the target population and failure to address possible confounding factors in making the conclusion that HCD helps make a usable product. This issue of the sampling methodology is also raised by the authors of the included studies; future HCD research should note this point in their strategic planning accordingly.

Finally, the authors stress that the aim of this systematic review was not to assess the implementation of HCD in creating mHealth for older adults or the effectiveness of mHealth interventions. The objective was to explore existing literature and establish recommendations and pitfalls for subsequent HCD projects. The older adults might be a narrow target population, but being the more sensitive and vulnerable group, the insight gained could be applicable to a wider range of users and help make future mHealth solutions more inclusive as well.

Conclusions

This review concludes that HCD can be used to create mHealth solutions for older adults and has summarized the process based on the 4 HCD steps with additional recommendations. The findings of this review can help designers, developers, and researchers gain an overview of HCD for older adults and implement the framework in their projects. The growing body of literature is encouraging, but more evidence-based results of HCD on creating mHealth for older adults are still needed. Future research should also focus on applying artificial intelligence and machine learning in HCD and utilizing the framework to create novel mHealth solutions for the population.

Acknowledgments

This research project was funded by the Ratchadapisek Sompoch Endowment Fund (2021) under Telehealth Cluster, Chulalongkorn University. The authors also thank the Center of Excellence in Hepatitis and Liver Cancer, Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, and Thailand Research Fund (RTA6280004).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2009 checklist of this systematic review.
[PDF File (Adobe PDF File), 96 KB - mhealth_v10i1e29512_app1.pdf]

References

1. Cai RA, Beste D, Chaplin H, Varakliotis S, Suffield L, Josephs F, et al. Developing and Evaluating JIApp: Acceptability and Usability of a Smartphone App System to Improve Self-Management in Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth 2017 Aug 15;5(8):e121 [FREE Full text] [doi: 10.2196/mhealth.7229] [Medline: 28811270]
2. Cushing CC, Fedele DA, Patton SR, McQuaid EL, Smyth JM, Prabhakaran S, et al. Responsive Asthma Care for Teens (ReACT): development protocol for an adaptive mobile health intervention for adolescents with asthma. BMJ Open 2019 Aug 20;9(8):e030029 [FREE Full text] [doi: 10.1136/bmjopen-2019-030029] [Medline: 31434777]
3. Hamel LM, Thompson HS, Albrecht TL, Harper FW. Designing and Testing Apps to Support Patients With Cancer: Looking to Behavioral Science to Lead the Way. JMIR Cancer 2019 Apr 22;5(1):e12317 [FREE Full text] [doi: 10.2196/12317] [Medline: 31066691]
4. Baek H, Suh J, Kang S, Kang S, Lim TH, Hwang H, et al. Enhancing User Experience Through User Study: Design of an mHealth Tool for Self-Management and Care Engagement of Cardiovascular Disease Patients. JMIR Cardio 2018 Feb 09;2(1):e3 [FREE Full text] [doi: 10.2196/cardio.9000] [Medline: 31758783]
5. Sobrinho A, da Silva LD, Perkusich A, Pinheiro ME, Cunha P. Design and evaluation of a mobile application to assist the self-monitoring of the chronic kidney disease in developing countries. BMC Med Inform Decis Mak 2018 Jan 12;18(1):7 [FREE Full text] [doi: 10.1186/s12911-018-0587-9] [Medline: 29329530]
6. Mayberry LS, Berg CA, Harper KJ, Osborn CY. The Design, Usability, and Feasibility of a Family-Focused Diabetes Self-Care Support mHealth Intervention for Diverse, Low-Income Adults with Type 2 Diabetes. J Diabetes Res 2016;2016:7586385-7586313 [FREE Full text] [doi: 10.1155/2016/7586385] [Medline: 27891524]
7. Giunti G, Mylonopoulou V, Rivera Romero O. More Stamina, a Gamified mHealth Solution for Persons with Multiple Sclerosis: Research Through Design. JMIR Mhealth Uhealth 2018 Mar 02;6(3):e51 [FREE Full text] [doi: 10.2196/mhealth.9437] [Medline: 29500159]

XSL•FO
RenderX

8. Fortuna KL, Lohman MC, Gill LE, Bruce ML, Bartels SJ. Adapting a Psychosocial Intervention for Smartphone Delivery to Middle-Aged and Older Adults with Serious Mental Illness. Am J Geriatr Psychiatry 2017 Aug;25(8):819-828 [FREE Full text] [doi: 10.1016/j.jagp.2016.12.007] [Medline: 28169129]
9. mHealth: New Horizons for Health Through Mobile Technologies. Geneva, Switzerland: WHO Press; 2011.
10. Ryu S. Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth 2009 (Global Observatory for eHealth Series, Volume 2). Geneva, Switzerland: WHO Press; 2012:153-159.
11. Schwamm LH, Erskine A, Licurse A. A digital embrace to blunt the curve of COVID19 pandemic. NPJ Digit Med 2020 May 4;3(1):64 [FREE Full text] [doi: 10.1038/s41746-020-0279-6] [Medline: 32377575]
12. Caulfield BM, Donnelly SC. What is Connected Health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
13. The mobile economy 2020. GSMA Intelligence. URL: https://data.gsmaintelligence.com/api-web/v2/research-file-download?id=51249388&file=2915-260220-Mobile-Economy.pdf [accessed 2020-11-12]
14. Gorski I, Bram JT, Sutermaster S, Eckman M, Mehta K. Value propositions of mHealth projects. J Med Eng Technol 2016 Aug 12;40(7-8):400-421. [doi: 10.1080/03091902.2016.1213907] [Medline: 27687907]
15. Wu X, Guo X, Zhang Z. The Efficacy of Mobile Phone Apps for Lifestyle Modification in Diabetes: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
16. Li R, Liang N, Bu F, Hesketh T. The Effectiveness of Self-Management of Hypertension in Adults Using Mobile Health: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: 10.2196/17776] [Medline: 32217503]
17. Ramsey RR, Plevinsky JM, Kollin SR, Gibler RC, Guilbert TW, Hommel KA. Systematic Review of Digital Interventions for Pediatric Asthma Management. J Allergy Clin Immunol Pract 2020 Apr;8(4):1284-1293 [FREE Full text] [doi: 10.1016/j.jaip.2019.12.013] [Medline: 31870809]
18. Sundin P, Callan J, Mehta K. Why do entrepreneurial mHealth ventures in the developing world fail to scale? J Med Eng Technol 2016 Aug 09;40(7-8):444-457. [doi: 10.1080/03091902.2016.1213901] [Medline: 27686003]
19. Desveaux L, Soobiah C, Bhatia RS, Shaw J. Identifying and Overcoming Policy-Level Barriers to the Implementation of Digital Health Innovation: Qualitative Study. J Med Internet Res 2019 Dec 20;21(12):e14994 [FREE Full text] [doi: 10.2196/14994] [Medline: 31859679]
20. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
21. Hilliard ME, Hahn A, Ridge AK, Eakin MN, Riekert KA. User Preferences and Design Recommendations for an mHealth App to Promote Cystic Fibrosis Self-Management. JMIR Mhealth Uhealth 2014 Oct 24;2(4):e44 [FREE Full text] [doi: 10.2196/mhealth.3599] [Medline: 25344616]
22. Van Velthoven MH, Cordon C. Sustainable Adoption of Digital Health Innovations: Perspectives From a Stakeholder Workshop. J Med Internet Res 2019 Mar 25;21(3):e11922 [FREE Full text] [doi: 10.2196/11922] [Medline: 30907734]
23. United Nations, Department of Economic and Social Affairs, Population Division. World Population Aging 2019 - Highlights (ST/ESA/SER.A/430). Geneva, Switzerland: United Nations; 2019.
24. Salive ME. Multimorbidity in older adults. Epidemiol Rev 2013 Jan 31;35(1):75-83. [doi: 10.1093/epirev/mxs009] [Medline: 23372025]
25. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
26. ISO 9241-210: Ergonomics of Human-System Interaction - Part 210: Human-Centered Design for Interactive Systems. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2020-10-12]
27. Siddaway AP, Wood AM, Hedges LV. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. Annu Rev Psychol 2019 Jan 04;70:747-770. [doi: 10.1146/annurev-psych-010418-102803] [Medline: 30089228]
28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
29. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018 Dec 18;34(4):285-291. [doi: 10.3233/efi-180221]
30. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold Stories in User-Centered Design of Mobile Health: Practical Challenges and Strategies Learned From the Design and Evaluation of an App for Older Adults With Heart Failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
31. Cornet VP, Daley CN, Srinivas P, Holden RJ. User-Centered Evaluations with Older Adults: Testing the Usability of a Mobile Health System for Heart Failure Self-Management. Proc Hum Factors Ergon Soc Annu Meet 2017 Sep 28;61(1):6-10 [FREE Full text] [doi: 10.1177/1541931213601497] [Medline: 30930610]

XSL•FO
RenderX

32. Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, et al. Human-Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users. JMIR Mhealth Uhealth 2017 May 30;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
33. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A Human-Centered Design Methodology to Enhance the Usability, Human Factors, and User Experience of Connected Health Systems: A Three-Phase Methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
34. Petersen CL, Halter R, Kotz D, Loeb L, Cook S, Pidgeon D, et al. Using Natural Language Processing and Sentiment Analysis to Augment Traditional User-Centered Design: Development and Usability Study. JMIR Mhealth Uhealth 2020 Aug 07;8(8):e16862 [FREE Full text] [doi: 10.2196/16862] [Medline: 32540843]
35. Srinivas P, Cornet V, Holden R. Human factors analysis, design, and evaluation of Engage, a consumer health IT application for geriatric heart failure self-care. Int J Hum Comput Interact 2017 Dec 02;33(4):298-312 [FREE Full text] [doi: 10.1080/10447318.2016.1265784] [Medline: 30429638]
36. Stara V, Harte R, Di Rosa M, Rossi L, ÓLaighin G. Toward a connected health system for older adults: lessons learned. In: Goonetilleke R, Karwowski W, editors. Advances in Physical Ergonomics and Human Factors. Cham, Switzerland: Springer; 2016.
37. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product From the ESRC Methods Programme. 2006. URL: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf [accessed 2022-01-10]
38. Minto B. The Minto Pyramid Principle: Logic in Writing, Thinking and Problem Solving (Reprinted Edition). London, UK: Minto Books International, Inc; 2012.
39. Hong QN, Pluye P, Fabregues S, Bartlett G, Boardman F, Cargo M, et al. Reporting the Results of the MMAT (Version 2018). 2018. URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/140056890/Reporting%20the%20results%20of%20the%20MMAT.pdf [accessed 2020-10-12]
40. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1990 Presented at: SIGCHI Conference on Human Factors in Computing Systems; March 1990; Seattle, WA p. 249. [doi: 10.1145/97243.97281]
41. Nielsen J. Usability Engineering. San Francisco, CA: Morgan Kaufmann Publishers Inc; 1994.
42. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McLelland I, Weerdmeester BA, editors. Usability Evaluation in Industry. London, UK: Taylor and Francis; 1996:189-194.
43. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances in Psychology 1988;52:139-183. [doi: 10.1016/s0166-4115(08)62386-9]
44. Holden RJ, Bodke K, Tambe R, Comer RS, Clark DO, Boustani M. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care 2016 Jul 22;5(1):25-27. [doi: 10.1177/2327857916051003]
45. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. Journal of Usability Studies 2009 Nov;5(1):34-45 [FREE Full text]
46. Lewis JR. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2016 Aug 06 Presented at: Human Factors and Ergonomics Society Annual Meeting; October 12-16, 1992; Santa Monica, CA p. 1259-1260. [doi: 10.1177/154193129203601617]
47. Lund AM. Measuring usability with the USE questionnaire. Usability Interface 2001;8(2):3-6 [FREE Full text]
48. Khajouei R, Peute L, Hasman A, Jaspers M. Classification and prioritization of usability problems using an augmented classification scheme. J Biomed Inform 2011 Dec;44(6):948-957 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.002] [Medline: 21782036]
49. Martínez-Pérez B, de la Torre-Díez I, Candelas-Plasencia S, López-Coronado M. Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications. J Med Syst 2013 Oct 7;37(5):9976. [doi: 10.1007/s10916-013-9976-x] [Medline: 24014255]
50. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
51. Harte RP, Glynn LG, Broderick BJ, Rodriguez-Molinero A, Baker PMA, McGuiness B, et al. Human centred design considerations for connected health devices for the older adult. J Pers Med 2014 Jun 04;4(2):245-281 [FREE Full text] [doi: 10.3390/jpm4020245] [Medline: 25563225]
52. Li C, Neugroschl J, Zhu CW, Aloysi A, Schimming CA, Cai D, et al. Design Considerations for Mobile Health Applications Targeting Older Adults. JAD 2021 Jan 05;79(1):1-8. [doi: 10.3233/jad-200485]
53. Banos O, Villalonga C, Garcia R, Saez A, Damas M, Holgado-Terriza JA, et al. Design, implementation and validation of a novel open framework for agile development of mobile health applications. BioMed Eng OnLine 2015;14(Suppl 2):S6. [doi: 10.1186/1475-925x-14-s2-s6]
54. Stara V, Harte R, Di Rosa M, Glynn L, Casey M, Hayes P, et al. Does culture affect usability? A trans-European usability and user experience assessment of a falls-risk connected health system following a user-centred design methodology carried out in a single European country. Maturitas 2018 Aug;114:22-26. [doi: 10.1016/j.maturitas.2018.05.002] [Medline: 29907242]

XSL•FO
RenderX

55. Association for the Advancement of Medical Instrumentation. ANSI/AAMI HE75-2009: Human Factors Engineering-Design of Medical Devices. Arlington, VA: Association for the Advancement of Medical Instrumentation; 2009.
56. Mylonopoulou V, Weilenmann A, Torgersson O, Jungselius B. Searching for empathy: a swedish study on designing for seniors. 2020 Presented at: NordiCHI '20: Proceedings of the 11th Nordic Conference on Human-Computer Interaction; October 25-29, 2020; Tallinn, Estonia p. 1-10. [doi: 10.1145/3419249.3420125]
57. Berman MA, Guthrie NL, Edwards KL, Appelbaum KJ, Njike VY, Eisenberg DM, et al. Change in Glycemic Control With Use of a Digital Therapeutic in Adults With Type 2 Diabetes: Cohort Study. JMIR Diabetes 2018 Feb 14;3(1):e4 [FREE Full text] [doi: 10.2196/diabetes.9591] [Medline: 30291074]
58. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. J Family Med Prim Care 2020 May;9(5):2207-2213 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_105_20] [Medline: 32754475]

Abbreviations

ASQ: After Scenario Questionnaire
CIED: cardiac implantable electronic devices
HCD: human-centered design
HIPAA: Health Insurance Portability and Accountability Act of 1996
I-IMR: Integrated Illness Management and Recovery
ISO: International Organization for Standardization
mHealth: mobile health
MMAT: Mixed Methods Appraisal Tool
NASA-TLX: NASA-Task Load Index
NVS: Newest Vital Sign
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSSUQ: Post-Study System Usability Questionnaire
QUEST: the Quebec User Evaluation of Satisfaction with Assistive Technology questionnaire
SUS: System Usability Scale
USE: Usefulness, Satisfaction, and Ease of use
WIISEL: Wireless Insole for Independent and Safe Elderly Living

Edited by G Eysenbach, L Buis; submitted 09.04.21; peer-reviewed by V Mylonopoulou, A Perkusich, V Stara; comments to author 27.04.21; revised version received 24.08.21; accepted 23.11.21; published 14.01.22.

Please cite as:
Nimmanterdwong Z, Boonviriya S, Tangkijvanich P
Human-Centered Design of Mobile Health Apps for Older Adults: Systematic Review and Narrative Synthesis
JMIR Mhealth Uhealth 2022;10(1):e29512
URL: https://mhealth.jmir.org/2022/1/e29512
doi:10.2196/29512
PMID:35029535

©Zethapong Nimmanterdwong, Suchaya Boonviriya, Pisit Tangkijvanich. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 14.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

mHealth Solutions for Perinatal Mental Health: Scoping Review and Appraisal Following the mHealth Index and Navigation Database Framework

Benedetta Spadaro[1*], MPhil; Nayra A Martin-Key[1*], PhD; Erin Funnell[1], BSc; Sabine Bahn[1,2], MD, PhD

[1]Cambridge Centre for Neuropsychiatric Research, Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom

[2]Psyomics Ltd, Cambridge, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Sabine Bahn, MD, PhD
Cambridge Centre for Neuropsychiatric Research
Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom
Phone: 44 1223 334151
Email: sb209@cam.ac.uk

Abstract

Background: The ever-increasing pressure on health care systems has resulted in the underrecognition of perinatal mental disorders. Digital mental health tools such as apps could provide an option for accessible perinatal mental health screening and assessment. However, there is a lack of information regarding the availability and features of perinatal app options.

Objective: This study aims to evaluate the current state of diagnostic and screening apps for perinatal mental health available on the Google Play Store (Android) and Apple App Store (iOS) and to review their features following the mHealth Index and Navigation Database framework.

Methods: Following a scoping review approach, the Apple App Store and Google Play Store were systematically searched to identify perinatal mental health assessment apps. A total of 14 apps that met the inclusion criteria were downloaded and reviewed in a standardized manner using the mHealth Index and Navigation Database framework. The framework comprised 107 questions, allowing for a comprehensive assessment of app origin, functionality, engagement features, security, and clinical use.

Results: Most apps were developed by for-profit companies (n=10), followed by private individuals (n=2) and trusted health care companies (n=2). Out of the 14 apps, 3 were available only on Android devices, 4 were available only on iOS devices, and 7 were available on both platforms. Approximately one-third of the apps (n=5) had been updated within the last 180 days. A total of 12 apps offered the Edinburgh Postnatal Depression Scale in its original version or in rephrased versions. Engagement, input, and output features included reminder notifications, connections to therapists, and free writing features. A total of 6 apps offered psychoeducational information and references. Privacy policies were available for 11 of the 14 apps, with a median Flesch-Kincaid reading grade level of 12.3. One app claimed to be compliant with the Health Insurance Portability and Accountability Act standards and 2 apps claimed to be compliant with General Data Protection Regulation. Of the apps that could be accessed in full (n=10), all appeared to fulfill the claims stated in their description. Only 1 app referenced a relevant peer-reviewed study. All the apps provided a warning for use, highlighting that the mental health assessment result should not be interpreted as a diagnosis or as a substitute for medical care. Only 3 apps allowed users to export or email their mental health test results.

Conclusions: These results indicate that there are opportunities to improve perinatal mental health assessment apps. To this end, we recommend focusing on the development and validation of more comprehensive assessment tools, ensuring data protection and safety features are adequate for the intended app use, and improving data sharing features between users and health care professionals for timely support.

(JMIR Mhealth Uhealth 2022;10(1):e30724) doi:10.2196/30724

XSL•FO
RenderX

KEYWORDS

digital mental health; perinatal mental health; pregnancy; MIND; mobile phone

Introduction

Background

Perinatal mental health disorders are among the most common complications of childbearing. Recent systematic reviews have reported a prevalence of 17% for postnatal depression [1] and 15% to 20% and 10% for antenatal and postnatal anxiety disorders [2,3], respectively. These numbers are comparable with the prevalence rates of gestational diabetes (14.5%) [4] and preterm birth (11%) [5]. Critically, pregnancy can be a triggering event; 10% to 20% of pregnant individuals with bipolar disorder experience a relapse during or after pregnancy, often culminating in a severe mental illness episode (ie, postpartum psychosis and mania) requiring hospitalization [6-8]. In some individuals, postpartum psychosis can also be the first manifestation of bipolar disorder [9,10]. If left untreated, perinatal mental illnesses can lead to poorer maternal quality of life, emotional suffering, and an increased risk of suicide and infanticide. Mental health disorders throughout the perinatal period have also been found to diminish mother–infant bonding [11], impair breast feeding [12,13], and, in some cases, predict poor outcomes in social-emotional and cognitive development of children [11]. In addition to considerable maternal and infant morbidity, perinatal mental disorders carry substantial health and social care costs to society. For instance, in the United Kingdom alone, perinatal depression, anxiety, and psychosis cost approximately £8.10 (US $10.90) billion for each 1-year cohort of births, with £1.20 (US $1.60) billion falling directly on the National Health Service and social services [14]. In Australia, overall costs of perinatal mental illness were estimated to reach Aus $7.30 (US $ 5.30) billion, with Aus $643 (US $466) million loss in productivity and Aus $227 (US $164) million incurred in health costs within the first year of perinatal mental illness [15]. In the United States, it was projected that untreated perinatal mood and anxiety disorders costed US $14 billion for the 2017 birth cohort from conception to 5 years postpartum, with the average cost per affected mother–child averaging at US $31,800 [16].

The ever-increasing pressure on health care systems and lack of time and resources have resulted in a staggering underrecognition of postnatal depression and other perinatal disorders [17,18], with approximately 50% of cases of postnatal depression being undiagnosed [19]. This is likely because of an array of individual-level and organizational-level barriers, including negative attitudes and stigma regarding diagnosis; a lack of understanding of perinatal mental health disorders among pregnant individuals, their partners, and health care professionals; and fear of consequences [20]. Other challenges include cultural and language factors, resource fragmentation, and poor policy implementation [21].

In this regard, digital mental health services such as web-based assessments and apps could help alleviate some of the pressure put on in-person health care services and overcome barriers to help seeking, providing an alternative or complementary option for widespread perinatal mental health care provision. Indeed, mobile technology is rapidly expanding into the field of well-being and health care, with mobile health (mHealth) being among the fastest growing sectors with a compound annual growth rate of 32.5% [22]. Immediacy, accessibility, and affordability are among the potential benefits of using digital tools to identify perinatal mental health concerns. In the field of perinatal mental health, recent studies have highlighted that new mothers and those with postnatal depression are interested in using health apps [23], and recent studies have highlighted a growing research effort in the implementation of mHealth tools for psychoeducation and prevention of perinatal mental health concerns [24-26]. The COVID-19 pandemic has fueled interest in digital mental health solutions, as elevated levels of stress coupled with reduced in-person care prompted changes in mental health care provisions [26,27]. As a result, telemedicine services have been widely and successfully adopted in everyday perinatal care and mental health care [28-30] paving the way for the uptake of other digital health innovations such as apps.

Objectives

Importantly, there still exists a large gap between interest in mental health digital tools [25] and a comprehensive understanding of the scientific integrity, clinical validity, and features of digital assessment tools for mental health [31-34]. In fact, 1 of the top 10 research priorities recently identified by the James Lind Alliance Priority Setting Partnership for digital technology in mental health care is to identify the best methods to evaluate and endorse mental health apps [35]. To this end, the objectives of this study are to identify apps that offer mental health screening or assessments for perinatal mental health available on the Google Play Store (Android) and the Apple App Store (iOS) and to review their features, including accessibility, privacy and security, clinical evidence, engagement style, and interoperability. Available apps were assessed using the mHealth Index and Navigation Database (MIND) framework [36,37]. The framework was developed by Lagan et al [36,37] in collaboration with the app evaluation model of the American Psychiatric Association (APA) [38], reflecting consensus from various stakeholders such as service users, social workers, psychiatrists, and data scientists to derive a ready-to-use resource for patients and clinicians alike [36,38,39]. The initial 38 open questions from the APA model served as the basis for the development of 107 questions that required binary (yes or no) or numeric responses covering app origin, functionality, security, engagement features, and clinical use. In the MIND framework, broad open questions from the APA model, for example, “What are the main engagement styles of the app?” were operationalized into 11 different types of engagement features. Similarly, the APA question “Is there a transparent privacy policy that is clear and accessible before use?” was operationalized into 2 objective questions: one regarding the presence of the privacy policy requiring a binary (yes or no) answer and the other prompting the rater to measure the reading level of the privacy policy (numeric response) to evaluate clarity.

XSL•FO
RenderX

Thus, using defined and discrete evaluative questions, the MIND framework aims to be more objective and reproducible than the APA model.

The findings from this timely appraisal of perinatal diagnostic and screening apps have important implications for clinical practice and for the development of innovative ways to provide mental health care provisions throughout this complex time.

Methods

Overview

The objectives of this scoping review are to identify apps that offer perinatal mental health screening or assessments and to review their features against the MIND framework [36,37]. The scope comprised apps whose intended user populations specifically included adults in the perinatal period (ie, pregnant or had recently given birth) or health care professionals operating in perinatal health care. In consultation with a practicing psychiatrist (SB), interventions of interest included apps presenting questions and answers based digital screening and diagnostic tools completed by an individual or a health care professional on behalf of the individual, used for mental health screening, or as an aid in clinical decision-making.

We reported the review following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines (Multimedia Appendix 1) [40].

Search Strategy

Search terms to identify apps developed specifically for perinatal mental health were identified through a preliminary search of the Apple App Store and Google Play Store. Relevant synonyms and layperson alternatives were also included in the search. Layperson alternatives were included to capture app results as searched by consumers who may not use technical terminology. As a result, the following keyword combinations were used: *moms mental health*, *moms mental health screening*, *moms mental health assessment*, *mums mental health*, *mums mental health assessment*, *mums mental health screening*, *pregnancy mental health*, *maternal mental health*, *pregnancy mental health assessment*, *pregnancy mental health screening*, *perinatal*, *perinatal mental health*, *perinatal mental health assessment*, *perinatal mental health screening*, *postpartum*, *postpartum mental health*, *postpartum mental health assessment*, *mental health screening*, and *mental health assessment*.

These terms were used to search the 2 most widely used smartphone app stores, Apple App Store and Google Play Store, between January and February 2021, to identify publicly available apps.

Inclusion Criteria

Apps were then shortlisted using the following inclusion criteria defined in consultation with a general adult psychiatrist (SB) and a practicing specialist perinatal psychiatrist:

1. Intended users of the app included at least 1 of the following groups: perinatal population and perinatal health care professionals.
2. The app offered a screening tool for any of the following conditions: bipolar disorder, major depressive disorder, postnatal depression, obsessive compulsive disorder, antenatal or postnatal anxiety, generalized anxiety disorder, agoraphobia, tokophobia, social phobia, panic disorder, insomnia, schizophrenia, (postpartum) psychosis, eating disorders (bulimia nervosa and anorexia nervosa), emotionally unstable personality disorder, alcohol abuse, substance abuse, (complex) posttraumatic stress disorder (PTSD), birth trauma, acute stress disorder, and adjustment disorder.
3. The app was available for download through the official Google Play or Apple App stores, and its description was available on the store.

Exclusion Criteria

Apps that were not intended specifically for use by the perinatal population or health care professionals operating in the perinatal health field were excluded. Apps presenting screening or assessment tools designed solely for any type of health care professional training and examination preparation purposes were also excluded.

Screening and App Selection

As the searches were performed on the Google Play Store and the Apple App Store, app names and links to the app stores were recorded on a Microsoft Excel spreadsheet. Duplicate apps retrieved using multiple search terms were then removed. A total of 2 independent reviewers (BS and NAMK) performed a blinded screening of the descriptions of all the identified apps. To decide whether the apps should be examined further, the independent reviewers assessed their eligibility against the inclusion criteria. Apps were labeled as *exclude*, *include*, or *maybe*. Any disagreements among the reviewers were discussed until a consensus was reached. The included apps were then analyzed and scored against the MIND framework [36].

Assessment

Apps meeting the inclusion criteria were downloaded onto either a Pixel 5 (Android version 11) or an iPhone X (iOS version 14.2) for complete assessment. One reviewer (BS) transferred the MIND framework questions from the supplementary materials in Lagan et al into Microsoft Excel spreadsheets for data extraction and assessment of app features. Two independent reviewers (BS and NAMK) extracted the relevant information on separate spreadsheets. The 2 reviewers (BS and NAMK) independently performed the assessment of each app following the MIND framework as outlined by Lagan et al [36]. The extracted data and assessment results were then compared, and any discrepancies or disagreements were resolved by consensus discussion between those 2 authors.

The MIND framework comprised 107 questions on features related to the following app characteristics: (1) app origin including developer characteristics; (2) app functionality, which includes platform, number of downloads, average user-scored star rating, need for network connectivity, language, and price; (3) inputs and outputs, such as the presence of surveys, reminders, access to camera and microphone; (4) privacy and security features, including presence of an accessible privacy

XSL•FO
RenderX

policy, data sharing policies and opt-outs, the presence of a crisis management feature; (5) evidence and clinical foundations, including adherence to app description claims, availability of evidence from feasibility studies, and compliance to clinical guidelines; (6) features and engagement style, which comprises characteristics such as the presence of tracking features, journaling, educational material, and peer support; (7) app use characteristics such as target audience and whether it is a self-managed tool or it is used together with a clinician; and (8) interoperability and data sharing matters, such as data ownership and interoperability with electronic medical records systems.

Reading level of the privacy policies of apps was calculated using an automatic text readability checker [41], as indicated in the MIND framework, resulting in a Flesh-Kincaid grade level [42], indicating a readability score corresponding to the US education grade level required for the reader to understand the text.

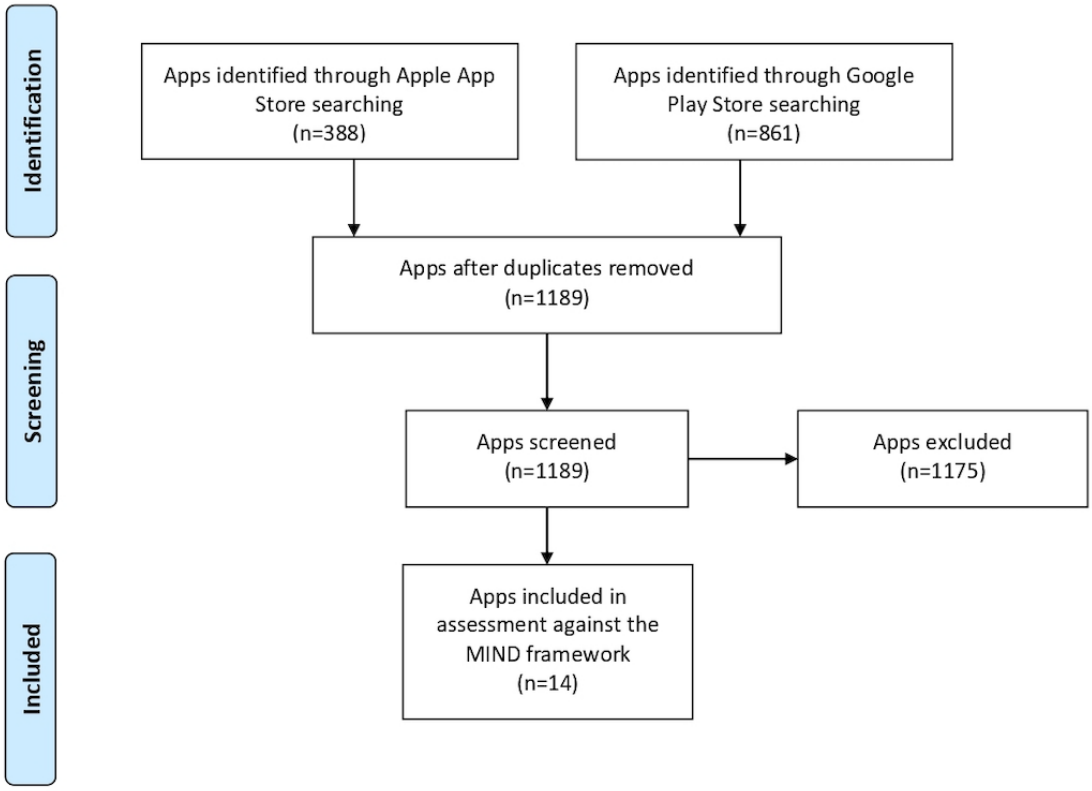
The MIND framework questions were designed to be answerable by any trained rater—clinician, peer, end user—and inform the identically titled public-facing database: the MIND [37,43]. The database was designed and implemented by the Division of Digital Psychiatry, a collaborative research group at the Beth Israel Deaconess Medical Center, a Harvard Medical School affiliate in Boston (MA, United States) that strives to create a comprehensive, easily searchable and updatable app database where apps are reviewed by trained raters following the MIND framework [36], and users can view app attributes and compare ratings. Hence, after performing the assessment of each of the apps following the MIND framework [36], we (BS and NAMK) searched the MIND database to compare our results with those of independent raters (Multimedia Appendix 2).

Results

Search Overview

After removing all duplicates, a total of 1189 unique apps were identified, with 801 apps from the Google Play Store and 388 apps from the Apple App Store. Duplicate apps retrieved using multiple search terms were removed. After reviewing the description of the apps, a total of 1175 apps were excluded as they were of no relevance (Figure 1 and Table 1). A total of 14 apps were included for the analysis and scored against the MIND framework [36] (Multimedia Appendix 3). Overall, only 4 apps could be partially assessed on the basis of the information extracted from the app description on the Google Play Store and Apple App Store. Of these, 2 required a referral by a health care provider, and 2 apps repeatedly crashed; hence, the content could not be assessed.

Figure 1. Flowchart of the search and selection strategy. MIND: mHealth Index and Navigation Database.



XSL•FO
RenderX

Table 1. Excluded apps' categories and frequencies.

App category	Category description	Frequencies,[a] n (%)
Mental health screening and tracking (not perinatal)	Mental health screening apps not targeted to the perinatal population and mood tracking apps.	95 (7.99)
Mental health information and interventions	Apps offering mindfulness, meditation, journaling, and sleep help, apps offering mental health coaching (eg, panic attack management and suicide prevention) as well as cognitive behavioral therapy features, and apps offering mental-health information and psychoeducation.	399 (33.56)
Perinatal physical health and well-being	Apps offering information about physical changes occurring over pregnancy and postpartum as well as information on breastfeeding and well-being during the perinatal period.	102 (8.58)
Physical health information and tracking	Apps providing information on neonatal care and physical health, female and reproductive health, general physical health, and health tracking features (eg, heart rate).	99 (8.33)
Educational and reference material for health care professionals	Apps providing revision material for medical school and nursing exams as well as information for health care professionals on care pathways, and apps for scientific conferences and journals.	120 (10.09)
Telemedicine	Apps acting as platforms for video consultations, remote patient monitoring and care.	43 (3.62)
Fitness	Apps offering workouts and, in some cases, nutritional advice for pregnant individuals and mothers, and also fitness apps targeted to the general population.	147 (12.36)
Other	Miscellaneous apps eg, apps for audiobooks, music, shopping, traveling, games, brain training or logic games, cuisine, social media, and personality tests	170 (14.30)

[a]Total sample size includes all unique apps identified in the search (N=1189).

App Origin and Functionality

Most apps were developed by for-profit providers (n=10) or private individuals (n=2), followed by trusted health care companies (n=2). Of the apps developed by trusted health care companies, 1 was developed in collaboration with an academic institution.

Of the 14 included apps, 3 were only available on Android devices, 4 were available only on Apple iOS devices, and 7 had versions for both platforms. Approximately one-third of the apps (n=5) had been updated within the last 180 days, with 11 apps updated within the last 13 months.

Most apps did not have sufficient reviews to display the average review rating. Of the apps that could be accessed, over half could work offline (n=7), not requiring an internet connection after downloading, and 2 apps could not function in the absence of an internet connection or required connectivity for some of their features. None of the apps presented accessibility features such as text size adjustment, a text-to-voice option, or color-blind color scheme options.

Examining costs, 8 apps were free, 1 was free to download, but it is unknown if it offered in-app purchases, 2 offered in-app purchases, or redirected the user to paid services for psychotherapy, whereas 2 apps were priced at US $0.99, and 1 was priced at US $1.09.

Inputs and Outputs

All the apps offered self-assessment questionnaires to screen for perinatal depression. A total of 12 apps offered the Edinburgh Postnatal Depression Scale (EPDS) in its original version [44] or in slightly adapted versions. In addition to the EPDS, 1 app included the Patient Health Questionnaire–8-item scale [45], the Generalized Anxiety Disorder 7-item scale [46], the Insomnia Severity Index [47,48], the Perceived Stress Index [49], and questions about PTSD from the Mini-International Neuropsychiatric Interview [50] with attribution to original sources. A total of 3 apps offered the EPDS as well as additional tests not specifically targeted to the perinatal population such as tests for anxiety disorders, attention deficit hyperactivity disorder, PTSD, alcohol and substance use disorders, other various forms of addiction, eating disorders, personality disorders, bipolar disorder, and schizophrenia.

Most of the apps (n=8) provided the user with mental health screening scores and short summaries of their clinical significance. In addition, 2 apps provided a graphical depiction of the assessment results over time. A total of 3 apps offered the option of receiving reminder notifications to prompt the user to complete mental health screenings regularly.

Upon informing the user of their test scores, 8 apps recommended discussing results with a health care professional, 1 displayed the EPDS score without duty-of-care messaging, and the remaining 5 apps could not be viewed in full; therefore, it could not be determined whether users were advised to consult a health care professional. A total of 5 apps displayed duty-of-care messages and provided the user with crisis hotlines; in 4 cases, these were US-based hotlines and in 1 case, these were UK-based hotline; therefore, crisis advice was not personalized to the location of the user. Only 3 apps instantly displayed a duty-of-care message referring the user to clinical support and crisis hotlines if symptoms regarding suicidal ideation or self-harm were endorsed.

In addition to the screening tools, 2 apps presented a free writing, diary-like feature, and 1 app allowed the user to add notes after having taken the screening test. A total of 6 apps offered psychoeducation information or references to psychoeducational materials. An additional app presented psychoeducation videos that could not be viewed, and another app reported a link to psychoeducation information that was not functioning.

XSL•FO
RenderX

Features and Engagement Style

Aside from screeners, assessments, reminders, and basic psychoeducation information, the apps did not include the great majority of engagement features listed in the MIND framework [36]. One app offered the user the possibility to connect with therapists online, whereas another app redirected users to BetterHelp, a web-based portal providing mental health services. One app allowed users to join a community on Facebook to share experiences and advice on mental health concerns. An additional app had a *safety plan* feature, allowing users to save motivational sentences and the details of up to 5 contacts to call in case of suicidal thoughts. None of the apps presented gamification features, such as gaining points or prizes for completing mental health assessments.

Privacy and Security

Privacy policies were available either as a link from the app store description or in the app for 11 of the 14 apps. Overall, 1 app claimed to be compliant with the Health Insurance Portability and Accountability Act standards, and 2 additional apps claimed to be compliant with General Data Protection Regulation. Data use and purpose was declared in all the available privacy policies: in 2 cases data use was not detailed and in a further case, the information seemed to be related to the website of the developer and not specific to the app. A description of measures aimed at secure data collection and sharing was present in 6 of the 11 policies available. A total of 5 apps stated in their privacy policies that personal health information (PHI), including name, birthday, and mental health information would not leave the app, whereas PHI was shared in the other 6 apps. One-third of the apps in which PHI was shared did not report measures aimed at secure data collection and sharing. Deidentified data were shared by 7 apps, 5 of which also shared anonymized or aggregated user data. Only 2 apps of those that collected and shared data specifically stated in their privacy policies that the user could opt out of data collection. Three apps offered the option to delete all data related to the user upon request, whereas 2 apps allowed only for partial deletion of personal data. Out of the 11 privacy policies considered, 6 mentioned a crisis management feature (eg, a hotline number was included at the end of the mental health assessment). However, as discussed above, only 3 apps presented the user with an instant duty-of-care message upon presentation of suicidal ideation or self-harm. Finally, the median reading level of the privacy policies as measured using the Flesch-Kincaid reading grade level was 12.3, which corresponds to 12th grade, the final year of secondary school in the United States.

Evidence, Clinical Foundations, Use, and Interoperability

Of the apps that could be accessed (n=10), all appeared to fulfill the claims stated in their descriptions. Importantly, the remaining 4 apps could not be evaluated on this criterion because they either crashed upon app launch or could not be accessed. All the assessed apps were patient-facing, with 3 of them being designed for use by both clinicians and patients.

One app referenced a relevant study conducted to test the app [51], where the usability of the app as a screening and management tool for perinatal depression was explored by gathering feedback from women in interviews. However, the study did not assess the efficacy of the diagnostic tool or psychoeducation content. One additional app referred to published, peer-reviewed studies, but it was unclear if the app tested in those studies corresponded to the current app version and the studies were not performed in perinatal populations.

All the apps provided a warning for use, highlighting that the mental health assessment result should not be interpreted as a diagnosis and that the app was not a substitute for medical care.

As a result, the apps were all regarded as *reference apps* and not *as self-help* tools. In the MIND framework, self-help apps are defined as providing activities that can be used for self-help and self-management, such as mood or symptom tracking or mindfulness exercises, whereas reference apps are defined as providing information and references but not necessarily activities. Although completing self-assessment questionnaires is a valuable activity, none of the apps that could be viewed in full offered activities that could help the user to manage the mental health concerns identified through the self-assessment activity.

The 3 apps allowed users to export or email their mental health test results. None of the apps seemed to have the necessary interoperability features to allow the sharing of app-gathered data to a medical record. The only app that offered an in-built connection with mental health therapists also allowed users to share their data with the therapist only after booking a therapy session.

Discussion

Principal Findings

Early mental health assessment strategies hold promise in supporting pregnant individuals and new mothers, but strategies vary widely among countries, with screening recommendations being subject to debates [52-54] and systematic reviews attempting to collect evidence to inform policy makers [55]. To date, access to mental health care is restricted to only a small proportion of pregnant individuals in need of mental health support [56]. During the COVID-19 pandemic, the increase in mental health concerns in the perinatal population has urged health care systems to expand care modalities [27,57]. As a result, guidelines compiled for perinatal care during the pandemic recommended asking pregnant individuals about their mental health at every antenatal and postnatal appointment and encouraged the use of digital means to deliver support [58].

Telemedicine has been the main digital tool used by health care professionals transitioning to remote care models, but apps have also started to be featured in mental health programs offered by health care providers and universities [59,60]. In a recent study conducted by our group, women, partners, and midwives expressed a strong interest in using a digital mental health assessment to screen, diagnose, and triage perinatal mental concerns [26]. This finding resonates with previous evidence showing that pregnant individuals are increasingly using digital

XSL•FO
RenderX

tools as a source of health information during pregnancy and to enhance their understanding and involvement in pregnancy-related decision-making [61]. Moreover, studies also support the view that individuals are open to mental health discussions at perinatal visits and that delivery modality (eg, paper vs tablet) does not affect acceptability [62,63].

The interest in perinatal digital tools set a solid stepping stone for apps looking to meet the growing demand for accessible mental health screens and assessments. However, in contrast to this positive outlook, our results revealed an unsatisfactory landscape of existing app options for perinatal mental health screening and assessment. First, 14.3% (170/1189) of the results from the keyword search were completely unrelated to mental health and the perinatal period (*Other*; Table 1), which may be disheartening for users looking for help and relevant tools. Moreover, 12.36% (147/1189) of the excluded apps offered fitness programs, including workouts and weight loss (*Fitness*; Table 1), with such a focus on body image being potentially deleterious and triggering in the perinatal mental health context. A similar prevalence of apps targeting physical appearance and fitness was also reported by a recent systematic review specific to mHealth interventions for peripartum mood disorders [64].

In summary, of the included apps, several had not been recently updated, lacked accessibility features (eg, text size adjustments, text-to-voice options), and often presented functionality issues that could be deleterious and disheartening to users referring to the app for support. The screening tool most frequently encountered in this review was the EPDS. The EPDS can be used to screen for depression and anxiety [65]. However, most app descriptions and EPDS results reports focused heavily on perinatal depression. Only 1 app presented a more comprehensive screening pathway using questions and validated screening tools for depression, anxiety, insomnia, and PTSD symptoms. None of the apps acted as a diagnostic tool, instead they acted as screening tools. Indeed, the tools used by the apps such as the EPDS, Generalized Anxiety Disorder 7-item scale, Insomnia Severity Index, Patient Health Questionnaire–8-item scale, and Perceived Stress Index are screening tools aimed at identifying individuals who may benefit from further assessment. Screening tools are designed to have high sensitivity, whereas diagnostic tools are designed to have good content validity, test–retest reliability, good interrater reliability, and high specificity [66]. Using self-report inventories designed for screening purposes as a single means of deriving a diagnosis is inadequate and must be avoided. An in-depth interview with a clinician following the diagnostic criteria defined by the Diagnostic and Statistical Manual of Mental Disorders, fifth edition or the International Classification of Diseases, eleventh edition remains the gold standard for diagnostic assessment. Therefore, apps using self-reported inventories are far from being comprehensive enough to act as diagnostic tools, whereas they may be useful as a first-step screening tool that alerts individuals of the need for further assessment.

Indeed, if the screening result is positive, guiding app users toward a full diagnostic assessment is critical. Our review highlighted that, in some cases, the apps failed to provide clear, instant, and geographically relevant duty-of-care messages to users, even upon disclosure of self-harm or suicidality. Compliance with data security statutes and regulations (eg, Health Insurance Portability and Accountability Act and General Data Protection Regulation) has rarely been mentioned, and with a median Flesch-Kincaid reading grade level of 12.3, privacy policies were often above the suggested readability grade level (9-10) for the public [67]. Importantly, sharing of results with clinicians was enabled only by 3 apps, which allowed users to export or email their screening reports.

Strengths and Limitations

The search strategy used to identify perinatal mental health assessment apps was comprehensive, and a completely blinded dual review process was used to ensure all relevant apps were included and to decrease the risk of reviewer bias.

The assessment was conducted using the MIND framework developed in collaboration with the APA [36]. In a recent review, the MIND framework operationalized by Lagan et al [37] showed significant overlap with other 70 app evaluation frameworks, highlighting its comprehensiveness and flexibility as a consolidated framework for app assessments. However, it should be noted that the framework does not include questions about ease of use, visual appeal, layout, and graphics. These are arguably subjective app features and are probably best assessed within the relevant clinical context and user population.

Critically, assessment of the apps strongly depended on the information disclosed by app developers in the app description, privacy policies, and functionality of the app itself. Our search did not include web or proprietary digital platforms that are not commercially available as apps on consumer-facing platforms. Recent literature shows that proprietary digital tools options exist and are offered by perinatal care providers [68], and there is also a disconnect between commercially available apps and academically available apps for perinatal mental health, which have been reviewed elsewhere [64].

Unlike electronic journal databases, app stores are not designed for systematic search and export of data [69]. For instance, there are challenges in removing duplicates directly in app store searches, and search results may be inconsistent given the nature of search algorithms and personalized app content of commercial app stores. Thus, it is challenging to replicate the search strategy reliably [70]. Currently, no guidelines exist for the conduct and reporting of systematic searches of app stores, but efforts are being made to reach a consensus [70].

Finally, our analysis aimed to analyze apps specifically designed and targeted to the perinatal population. However, our search revealed that there is a plethora of mental health screening, tracking, and interventions that do not specifically target the perinatal population (Table 1). The use of such tools may be helpful in identifying underlying mental health conditions but may not be able to capture dimensions that are specific to the perinatal period. Moreover, cutoffs, specificity, and sensitivity parameters of tools strongly depend on the population used for validation. Thus, the reliability of screening tools in a setting or population different from that in which the tool was developed cannot be guaranteed [71].

XSL•FO
RenderX

Conclusions

This review of apps for perinatal mental health assessment following the MIND framework supports the view that there are gaps in the current app space. As a result, we recommend 3 areas of focus for app developers and clinicians in designing and evaluating apps for perinatal mental health assessments.

First, development of more comprehensive digital screening tools is required. Critically, although the identification of perinatal depression is facilitated by validated questionnaires such as the EPDS, there is less consensus on screening tools for other disorders with significant prevalence, including perinatal anxiety and substance use disorder. Systematic reviews of mental health screening tools [72-74] may be used to inform the development and assessment of digital questionnaires that aim to provide a more comprehensive screening. Formal validation against a gold standard is then required to establish the reliability and accuracy parameters of screening tools in the setting and population of interest.

Second, the importance of safety features cannot be overstated. Any developer aiming to design a mental health screening tool should be responsible for keeping PHI safe and providing adequate information about sources of help in case of disclosure of self-harm or suicidality. To this end, following data security statutes and keeping apps up-to-date with geographically relevant sources of help and crisis hotlines represent the very first steps toward a more secure and responsible development of mental health apps.

Third, app developers and clinicians should strive to increase interoperability and data sharing. Digital self-reported screening tools can increase access to mental health support and aid triage only if the user is encouraged to share their screening results with a health care professional. To this end, data sharing must be easy for users and clinicians alike. With only 3 apps of the reviewed apps allowing to export screening results in some form (email or PDF), we believe there exists an opportunity to develop tools that better integrate with current medical records systems. Enhancing data sharing in a secure manner is likely to increase the use of mental health screening apps and contribute to a better therapeutic alliance between app users, developers, and clinicians.

Acknowledgments

This study was funded by the Stanley Medical Research Institute (grant number 07R-1888). We are grateful to perinatal psychiatrist Dr Nikolett Kabacs for the valuable input.

Conflicts of Interest

SB is a director of Psynova Neurotech Ltd and Psyomics Ltd and has financial interests in Psyomics Ltd. The other authors have no conflicts to declare.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[DOC File , 123 KB - mhealth_v10i1e30724_app1.doc]

Multimedia Appendix 2
Comparison of authors' assessment with ratings available in the mHealth Index and Navigation Database sorted by rating category (platform, developer type, etc).
[DOC File , 183 KB - mhealth_v10i1e30724_app2.doc]

Multimedia Appendix 3
Results of the assessment of the included apps (n=14) following the questions (n=107) of the mHealth Index and Navigation Database framework.
[XLSX File (Microsoft Excel File), 28 KB - mhealth_v10i1e30724_app3.xlsx]

References

1. Shorey S, Chee CY, Ng ED, Chan YH, Tam WW, Chong YS. Prevalence and incidence of postpartum depression among healthy mothers: a systematic review and meta-analysis. J Psychiatr Res 2018 Sep;104:235-248. [doi: 10.1016/j.jpsychires.2018.08.001] [Medline: 30114665]
2. Fawcett EJ, Fairbrother N, Cox ML, White IR, Fawcett JM. The prevalence of anxiety disorders during pregnancy and the postpartum period: a multivariate bayesian meta-analysis. J Clin Psychiatry 2019 Jul 23;80(4):18r12527 [FREE Full text] [doi: 10.4088/JCP.18r12527] [Medline: 31347796]
3. Micali N, Treasure J, Simonoff E. Eating disorders symptoms in pregnancy: a longitudinal study of women with recent and past eating disorders and obesity. J Psychosom Res 2007 Sep;63(3):297-303. [doi: 10.1016/j.jpsychores.2007.05.003] [Medline: 17719368]
4. Egan A, Dunne F. Epidemiology of gestational and pregestational diabetes mellitus. In: Lapolla A, Metzger BE, editors. Gestational Diabetes. A Decade After the HAPO Study. Basel: Karger; 2020:1-10.

XSL•FO
RenderX

5. Walani SR. Global burden of preterm birth. Int J Gynaecol Obstet 2020 Jul 10;150(1):31-33. [doi: 10.1002/ijgo.13195] [Medline: 32524596]
6. Stevens AW, Goossens PJ, van der Klein EA, Draisma S, Honig A, Kupka RW. Risk of recurrence of mood disorders during pregnancy and the impact of medication: a systematic review. J Affect Disord 2019 Apr 15;249:96-103. [doi: 10.1016/j.jad.2019.02.018] [Medline: 30769297]
7. Wesseloo R, Kamperman AM, Munk-Olsen T, Pop VJ, Kushner SA, Bergink V. Risk of postpartum relapse in bipolar disorder and postpartum psychosis: a systematic review and meta-analysis. Am J Psychiatry 2016 Feb 01;173(2):117-127. [doi: 10.1176/appi.ajp.2015.15010124] [Medline: 26514657]
8. Taylor CL, Broadbent M, Khondoker M, Stewart RJ, Howard LM. Predictors of severe relapse in pregnant women with psychotic or bipolar disorders. J Psychiatr Res 2018 Sep;104:100-107 [FREE Full text] [doi: 10.1016/j.jpsychires.2018.06.019] [Medline: 30015264]
9. Sharma V, Bergink V, Berk M, Chandra PS, Munk-Olsen T, Viguera AC, et al. Childbirth and prevention of bipolar disorder: an opportunity for change. Lancet Psychiatry 2019 Sep;6(9):786-792. [doi: 10.1016/S2215-0366(19)30092-6] [Medline: 30981755]
10. Jones I, Chandra PS, Dazzan P, Howard LM. Bipolar disorder, affective psychosis, and schizophrenia in pregnancy and the post-partum period. Lancet 2014 Nov 15;384(9956):1789-1799. [doi: 10.1016/S0140-6736(14)61278-2] [Medline: 25455249]
11. Nonnenmacher N, Noe D, Ehrenthal JC, Reck C. Postpartum bonding: the impact of maternal depression and adult attachment style. Arch Womens Ment Health 2016 Dec;19(5):927-935. [doi: 10.1007/s00737-016-0648-y] [Medline: 27296356]
12. Hoff CE, Movva N, Vollmar AK, Pérez-Escamilla R. Impact of maternal anxiety on breastfeeding outcomes: a systematic review. Adv Nutr 2019 Sep 01;10(5):816-826 [FREE Full text] [doi: 10.1093/advances/nmy132] [Medline: 31079143]
13. Figueiredo B, Canário C, Field T. Breastfeeding is negatively affected by prenatal depression and reduces postpartum depression. Psychol Med 2014 Apr;44(5):927-936. [doi: 10.1017/S0033291713001530] [Medline: 23822932]
14. Bauer A, Parsonage M, Knapp M, Lemmi V, Adelaja B. The costs of perinatal mental health problems. LSE Centre Mental Health. 2014. URL: https://www.centreformentalhealth.org.uk/sites/default/files/2018-09/costsofperinatalsummary.pdf [accessed 2021-12-27]
15. The cost of perinatal depression and anxiety in Australia. PwC Consulting Australia. 2019. URL: https://www.perinatalwellbeingcentre.org.au/Handlers/Download.ashx?IDMF=53aab8d3-c748-4818-abab-32a58d3c510f [accessed 2021-12-27]
16. Luca DL, Margiotta C, Staatz C, Garlow E, Christensen A, Zivin K. Financial toll of untreated perinatal mood and anxiety disorders among 2017 births in the United States. Am J Public Health 2020 Jun;110(6):888-896. [doi: 10.2105/AJPH.2020.305619] [Medline: 32298167]
17. Gavin N, Meltzer-Brody S, Glover V, Gaynes B. Is population-based identification of perinatal depression and anxiety desirable?: a public health perspective on the perinatal depression care continuum. In: Identifying Perinatal Depression and Anxiety. Hoboken, New Jersey, United States: John Wiley & Sons; 2014:11-31.
18. Ford E, Shakespeare J, Elias F, Ayers S. Recognition and management of perinatal depression and anxiety by general practitioners: a systematic review. Fam Pract 2017 Feb;34(1):11-19. [doi: 10.1093/fampra/cmw101] [Medline: 27660558]
19. Hewitt CE, Gilbody SM. Is it clinically and cost effective to screen for postnatal depression: a systematic review of controlled clinical trials and economic evidence. Br J Obstet Gynaecol 2009 Jul;116(8):1019-1027. [doi: 10.1111/j.1471-0528.2009.02148.x] [Medline: 19438499]
20. Smith M, Lawrence V, Sadler E, Easter A. Barriers to accessing mental health services for women with perinatal mental illness: systematic review and meta-synthesis of qualitative studies in the UK. BMJ Open 2019 Jan 24;9(1):e024803 [FREE Full text] [doi: 10.1136/bmjopen-2018-024803] [Medline: 30679296]
21. Bayrampour H, Hapsari AP, Pavlovic J. Barriers to addressing perinatal mental health issues in midwifery settings. Midwifery 2018 Apr;59:47-58. [doi: 10.1016/j.midw.2017.12.020] [Medline: 29353691]
22. mHealth economics 2016 - current status and trends of the mHealth app market. Research 2 Guidance. 2016. URL: https://research2guidance.com/product/mhealth-app-developer-economics-2016/ [accessed 2021-02-23]
23. Osma J, Barrera AZ, Ramphos E. Are pregnant and postpartum women interested in health-related apps? Implications for the prevention of perinatal depression. Cyberpsychol Behav Soc Netw 2016 Jun;19(6):412-415. [doi: 10.1089/cyber.2015.0549] [Medline: 27327069]
24. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
25. Sprenger M, Mettler T, Osma J. Health professionals' perspective on the promotion of e-mental health apps in the context of maternal depression. PLoS One 2017;12(7):e0180867 [FREE Full text] [doi: 10.1371/journal.pone.0180867] [Medline: 28704442]
26. Martin-Key NA, Spadaro B, Schei TS, Bahn S. Proof-of-concept support for the development and implementation of a digital assessment for perinatal mental health: mixed methods study. J Med Internet Res 2021 Jun 04;23(6):e27132 [FREE Full text] [doi: 10.2196/27132] [Medline: 34033582]

XSL•FO
RenderX

27. Osborne LM, Kimmel MC, Surkan PJ. The crisis of perinatal mental health in the age of Covid-19. Matern Child Health J 2021 Mar;25(3):349-352 [FREE Full text] [doi: 10.1007/s10995-020-03114-y] [Medline: 33543374]
28. Examples of alternate or reduced prenatal care schedules. American College of Obstetrics and Gynecology. 2020. URL: https://www.acog.org/clinical-information/physician-faqs/-/media/287cefdb936e4cda99a683d3cd56dca1.ashx [accessed 2021-12-27]
29. Kimmel M. Maternal mental health matters. N C Med J 2020;81(1):45-50 [FREE Full text] [doi: 10.18043/ncm.81.1.45] [Medline: 31908334]
30. Torous J, Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
31. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. Mhealth 2018;4:6 [FREE Full text] [doi: 10.21037/mhealth.2018.03.02] [Medline: 29682510]
32. Powell AC, Torous J, Chan S, Raynor GS, Shwarts E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth 2016;4(1):e15 [FREE Full text] [doi: 10.2196/mhealth.5176] [Medline: 26863986]
33. Yasini M, Marchand G. Toward a use case based classification of mobile health applications. Stud Health Technol Inform 2015;210:175-179. [Medline: 25991125]
34. Carpenter-Song E, Noel VA, Acquilano SC, Drake RE. Real-world technology use among people with mental illnesses: qualitative study. JMIR Ment Health 2018 Nov 23;5(4):e10652 [FREE Full text] [doi: 10.2196/10652] [Medline: 30470681]
35. Hollis C, Sampson S, Simons L, Davies EB, Churchill R, Betton V, et al. Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. Lancet Psychiatry 2018 Oct;5(10):845-854. [doi: 10.1016/S2215-0366(18)30296-7] [Medline: 30170964]
36. Lagan S, Aquino P, Emerson MR, Fortuna K, Walker R, Torous J. Actionable health app evaluation: translating expert frameworks into objective metrics. NPJ Digit Med 2020;3:100 [FREE Full text] [doi: 10.1038/s41746-020-00312-4] [Medline: 32821855]
37. Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. BMJ Open 2021 Mar 19;11(3):e047001 [FREE Full text] [doi: 10.1136/bmjopen-2020-047001] [Medline: 33741674]
38. The app evaluation model. American Psychiatric Association. URL: https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/the-app-evaluation-model [accessed 2021-12-27]
39. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019 Jun;1(2):52-54 [FREE Full text] [doi: 10.1016/S2589-7500(19)30013-5] [Medline: 33323229]
40. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
41. Automatic Readability Checker, a free readability formula consensus calculator. Readability Formulas. URL: https://readabilityformulas.com/free-readability-formula-tests.php [accessed 2021-12-27]
42. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
43. M-health Index and Navigation Database (MIND). URL: https://apps.digitalpsych.org/ [accessed 2021-12-27]
44. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987 Jun;150:782-786. [doi: 10.1192/bjp.150.6.782] [Medline: 3651732]
45. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
46. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
47. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
48. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 01;34(5):601-608 [FREE Full text] [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
49. Hewitt PL, Flett GL, Mosher SW. The Perceived Stress Scale: factor structure and relation to depression symptoms in a psychiatric sample. J Psychopathol Behav Assess 1992 Sep;14(3):247-257. [doi: 10.1007/bf00962631]
50. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-34. [Medline: 9881538]
51. Dyurich A, Oliver M. Use of the veedamom electronic app as a pregnancy treatment companion. J Feminist Fam Ther 2020 Jul 17;32(1-2):38-56. [doi: 10.1080/08952833.2020.1793562]
52. Chaudron LH, Wisner KL. Perinatal depression screening: let's not throw the baby out with the bath water!. J Psychosom Res 2014 Jun;76(6):489-491. [doi: 10.1016/j.jpsychores.2014.03.011] [Medline: 24840147]

XSL•FO
RenderX

53. Gemmill AW. The long gestation of screening programmes for perinatal depressive disorders. J Psychosom Res 2014 Sep;77(3):242-243. [doi: 10.1016/j.jpsychores.2014.06.017] [Medline: 25149035]
54. Thombs BD, Saadat N, Riehm KE, Karter JM, Vaswani A, Andrews BK, et al. Consistency and sources of divergence in recommendations on screening with questionnaires for presently experienced health problems or symptoms: a comparison of recommendations from the Canadian Task Force on Preventive Health Care, UK National Screening Committee, and US Preventive Services Task Force. BMC Med 2017 Aug 09;15(1):150 [FREE Full text] [doi: 10.1186/s12916-017-0903-8] [Medline: 28789659]
55. Hamel C, Lang E, Morissette K, Beck A, Stevens A, Skidmore B, et al. Screening for depression in women during pregnancy or the first year postpartum and in the general adult population: a protocol for two systematic reviews to update a guideline of the Canadian Task Force on Preventive Health Care. Syst Rev 2019 Jan 19;8(1):27 [FREE Full text] [doi: 10.1186/s13643-018-0930-3] [Medline: 30660183]
56. Byatt N, Xu W, Levin LL, Simas TA. Perinatal depression care pathway for obstetric settings. Int Rev Psychiatry 2019 May;31(3):210-228. [doi: 10.1080/09540261.2018.1534725] [Medline: 30701995]
57. Lebel C, MacKinnon A, Bagshawe M, Tomfohr-Madsen L, Giesbrecht G. Elevated depression and anxiety symptoms among pregnant individuals during the COVID-19 pandemic. J Affect Disord 2020 Dec 01;277:5-13 [FREE Full text] [doi: 10.1016/j.jad.2020.07.126] [Medline: 32777604]
58. Poon LC, Yang H, Kapur A, Melamed N, Dao B, Divakar H, et al. Global interim guidance on coronavirus disease 2019 (COVID-19) during pregnancy and puerperium from FIGO and allied partners: information for healthcare professionals. Int J Gynaecol Obstet 2020 Jun;149(3):273-286. [doi: 10.1002/ijgo.13156] [Medline: 32248521]
59. Melcher J, Torous J. Smartphone apps for college mental health: a concern for privacy and quality of current offerings. Psychiatr Serv 2020 Nov 01;71(11):1114-1119. [doi: 10.1176/appi.ps.202000098] [Medline: 32664822]
60. Mordecai D, Histon T, Neuwirth E, Heisler WS, Kraft A, Bang Y, et al. How Kaiser permanente created a mental health and wellness digital ecosystem. NEJM Catalyst 2021 Jan;2(1). [doi: 10.1056/cat.20.0295]
61. Lagan BM, Sinclair M, Kernohan WG. Internet use in pregnancy informs women's decision making: a web-based survey. Birth 2010 Jun;37(2):106-115. [doi: 10.1111/j.1523-536X.2010.00390.x] [Medline: 20557533]
62. Kingston D, Biringer A, van Zanten S, Giallo R, McDonald S, MacQueen G, et al. Pregnant women's perceptions of the risks and benefits of disclosure during web-based mental health e-screening versus paper-based screening: randomized controlled trial. JMIR Ment Health 2017 Oct 20;4(4):e42 [FREE Full text] [doi: 10.2196/mental.6888] [Medline: 29054833]
63. Yapp E, Howard LM, Kadicheeni M, Telesia LA, Milgrom J, Trevillion K. A qualitative study of women's views on the acceptability of being asked about mental health problems at antenatal booking appointments. Midwifery 2019 Jul;74:126-133. [doi: 10.1016/j.midw.2019.03.021] [Medline: 30953968]
64. Feldman N, Back D, Boland R, Torous J. A systematic review of mHealth application interventions for peripartum mood disorders: trends and evidence in academia and industry. Arch Womens Ment Health 2021 Apr 30:1-12 [FREE Full text] [doi: 10.1007/s00737-021-01138-z] [Medline: 33929636]
65. Mazzoni SE, Bott NL, Hoffman MC. Screening for perinatal anxiety. Am J Obstet Gynecol 2021 Jun;224(6):628-629. [doi: 10.1016/j.ajog.2021.03.004] [Medline: 33705747]
66. Anderson J, Michalak E, Lam R. Depression in primary care: tools for screening, diagnosis, and measuring response to treatment. BC Med J 2002;44(8):415-419 [FREE Full text]
67. Krumay B, Klar J. Readability of privacy policies. In: Singhal A, Vaidya J, editors. Data and Applications Security and Privacy. Cham: Springer; 2020:388-399.
68. Highet N, Gamble J, Creedy D. Perinatal mental health and psychosocial risk screening in a community maternal and child health setting: evaluation of a digital platform. Prim Health Care Res Dev 2018 May 18;20:1-7. [doi: 10.1017/s1463423618000336]
69. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
70. Grainger R, Devan H, Sangelaji B, Hay-Smith J. Issues in reporting of systematic review methods in health app-focused reviews: a scoping review. Health Informatics J 2020 Sep 11:2930-2945 [FREE Full text] [doi: 10.1177/1460458220952917] [Medline: 32914696]
71. Fellmeth G, Harrison S, Opondo C, Nair M, Kurinczuk JJ, Alderdice F. Validated screening tools to identify common mental disorders in perinatal and postpartum women in India: a systematic review and meta-analysis. BMC Psychiatry 2021 Apr 20;21(1):200 [FREE Full text] [doi: 10.1186/s12888-021-03190-6] [Medline: 33879130]
72. Sinesi A, Maxwell M, O'Carroll R, Cheyne H. Anxiety scales used in pregnancy: systematic review. BJPsych Open 2019 Jan;5(1):e5 [FREE Full text] [doi: 10.1192/bjo.2018.75] [Medline: 30762504]
73. Levis B, Negeri Z, Sun Y, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) EPDS Group. Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) for screening to detect major depression among pregnant and postpartum women: systematic review and meta-analysis of individual participant data. Br Med J 2020 Nov 11;371:m4022 [FREE Full text] [doi: 10.1136/bmj.m4022] [Medline: 33177069]

XSL•FO
RenderX

74. Pritchett R, Kemp J, Wilson P, Minnis H, Bryce G, Gillberg C. Quick, simple measures of family relationships for use in clinical practice and research. A systematic review. Fam Pract 2011 Apr;28(2):172-187. [doi: 10.1093/fampra/cmq080] [Medline: 20978241]

Abbreviations

APA: American Psychiatric Association
EPDS: Edinburgh Postnatal Depression Scale
MIND: mHealth Index and Navigation Database
PHI: personal health information
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PTSD: posttraumatic stress disorder

Edited by L Buis; submitted 26.05.21; peer-reviewed by N Feldman, B Chaudhry, N Hussain-Shamsy; comments to author 01.08.21; revised version received 06.09.21; accepted 19.11.21; published 17.01.22.

Please cite as:
Spadaro B, Martin-Key NA, Funnell E, Bahn S
mHealth Solutions for Perinatal Mental Health: Scoping Review and Appraisal Following the mHealth Index and Navigation Database Framework
JMIR Mhealth Uhealth 2022;10(1):e30724
URL: https://mhealth.jmir.org/2022/1/e30724
doi:10.2196/30724
PMID:35037894

©Benedetta Spadaro, Nayra A Martin-Key, Erin Funnell, Sabine Bahn. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 17.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating Evidence-Based Content, Features of Exercise Instruction, and Expert Involvement in Physical Activity Apps for Pregnant Women: Systematic Search and Content Analysis

Melanie J Hayman[1*], PhD; Kristie-Lee Alfrey[1*], BPsychSc(Hons); Kim Waters[1*], BPsySc; Summer Cannon[1*], BPsychSc(Hons); Gregore I Mielke[2*], PhD; Shelley E Keating[3*], PhD; Gabriela P Mena[2*], PhD; Michelle F Mottola[4,5,6*], PhD; Kelly R Evenson[7*], PhD; Margie H Davenport[8*], PhD; S Ariel Barlow[9*], BS; Emily Budzynski-Seymour[10*], MSc; Natalie Comardelle[9*], BS; Madison Dickey[9*], MPH; Cheryce L Harrison[11*], PhD; Maryam Kebbe[9*], PhD; Trine Moholdt[12,13*], PhD; Lisa J Moran[11*], PhD; Taniya S Nagpal[14*], PhD; Stephanie Schoeppe[1*], PhD; Stephanie Alley[1*], PhD; Wendy J Brown[2*], PhD; Susan Williams[1*], PhD; Lisa Vincze[15,16*], PhD

[1]Appleton Institute, School of Health, Medical and Applied Sciences, Central Queensland University, Rockhampton, Australia

[2]Centre for Research on Exercise, Physical Activity and Health, School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, Australia

[3]School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, Australia

[4]R Samuel McLaughlin Foundation Exercise and Pregnancy Laboratory, School of Kinesiology, University of Western Ontario, London, ON, Canada

[5]Department of Anatomy & Cell Biology, University of Western Ontario, London, ON, Canada

[6]Children's Health Research Institute, University of Western Ontario, London, ON, Canada

[7]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, United States

[8]Faculty of Kinesiology, Sport, and Recreation, University of Alberta, Edmonton, AB, Canada

[9]Reproductive Endocrinology & Women's Health Laboratory, Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, LA, United States

[10]Faculty of Sport, Health and Social Sciences, Solent University, Southampton, United Kingdom

[11]Monash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

[12]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

[13]Women's Clinic, St. Olavs University Hospital, Trondheim, Norway

[14]Faculty of Applied Health Sciences, School of Kinesiology, Brock University, Niagara Region, ON, Canada

[15]School of Health Sciences and Social Work, Griffith University, Gold Coast, Australia

[16]Menzies Health Institute Queensland, Griffith Health Centre, Gold Coast, Australia

[*]all authors contributed equally

Corresponding Author:
Melanie J Hayman, PhD
Appleton Institute
School of Health, Medical and Applied Sciences
Central Queensland University
Bruce Highway
Building 7/LG.12, North Rockhampton Campus
Rockhampton, 4701
Australia
Phone: 61 749306912
Email: m.j.hayman@cqu.edu.au

Abstract

Background: Guidelines for physical activity and exercise during pregnancy recommend that all women without contraindications engage in regular physical activity to improve both their own health and the health of their baby. Many women are uncertain how to safely engage in physical activity and exercise during this life stage and are increasingly using mobile apps to access health-related information. However, the extent to which apps that provide physical activity and exercise advice align with current evidence-based pregnancy recommendations is unclear.

XSL•FO
RenderX

Objective: This study aims to conduct a systematic search and content analysis of apps that promote physical activity and exercise in pregnancy to examine the alignment of the content with current evidence-based recommendations; delivery, format, and features of physical activity and exercise instruction; and credentials of the app developers.

Methods: Systematic searches were conducted in the Australian App Store and Google Play Store in October 2020. Apps were identified using combinations of search terms relevant to pregnancy and exercise or physical activity and screened for inclusion (with a primary focus on physical activity and exercise during pregnancy, free to download or did not require immediate paid subscription, and an average user rating of ≥4 out of 5). Apps were then independently reviewed using an author-designed extraction tool.

Results: Overall, 27 apps were included in this review (Google Play Store: 16/27, 59%, and App Store: 11/27, 41%). Two-thirds of the apps provided some information relating to the frequency, intensity, time, and type principles of exercise; only 11% (3/27) provided this information in line with current evidence-based guidelines. Approximately one-third of the apps provided information about contraindications to exercise during pregnancy and referenced the supporting evidence. None of the apps actively engaged in screening for potential contraindications. Only 15% (4/27) of the apps collected information about the user's current exercise behaviors, 11% (3/27) allowed users to personalize features relating to their exercise preferences, and a little more than one-third provided information about developer credentials.

Conclusions: Few exercise apps designed for pregnancy aligned with current evidence-based physical activity guidelines. None of the apps screened users for contraindications to physical activity and exercise during pregnancy, and most lacked appropriate personalization features to account for an individual's characteristics. Few involved qualified experts during the development of the app. There is a need to improve the quality of apps that promote exercise in pregnancy to ensure that women are appropriately supported to engage in exercise and the potential risk of injury, complications, and adverse pregnancy outcomes for both mother and child is minimized. This could be done by providing expert guidance that aligns with current recommendations, introducing screening measures and features that enable personalization and tailoring to individual users, or by developing a recognized system for regulating apps.

(JMIR Mhealth Uhealth 2022;10(1):e31607) doi:10.2196/31607

KEYWORDS

apps; exercise; mobile health; mHealth; mobile phone; physical activity; pregnancy; exercise prescription; evidence-based guidelines; app development; systematic review; mobile phone

Introduction

Background

Physical activity during pregnancy promotes maternal, fetal, and neonatal health [1]. The health benefits of prenatal physical activity include reduced risk of excessive gestational weight gain, gestational diabetes, pre-eclampsia, delivery complications, preterm birth, newborn complications, and postpartum depression [1]. As such, guidelines for physical activity and exercise during pregnancy recommend that all pregnant women without contraindications (in which the benefits of physical activity and exercise may be outweighed by risks associated with a medical condition) should undertake regular physical activity comprising at least 150 minutes of moderate- to vigorous-intensity aerobic activity each week, along with the incorporation of regular muscle-strengthening exercises (including pelvic floor exercises) [1-3]. Guidelines also identify safety considerations for physical activity and exercise in pregnancy, including absolute and relative contraindications to commencing (previously inactive women) or continuing (previously active women) activity, warning signs and symptoms to stop activity, and exercises to avoid [1,2].

Despite these recommendations and increased interest in health behaviors during pregnancy [4-6], few pregnant women achieve adequate physical activity and exercise [7-10]. A cohort study involving 3482 Norwegian women reported that only 14.6% of the pregnant women followed current guidelines for physical activity during pregnancy at 17-21 weeks' gestation [7]. One reason for such low adherence rates may be that women are uncertain how to engage safely in physical activity and exercise during this life stage [11-13]. Furthermore, women may receive limited or inaccurate advice on physical activity and exercise participation from health care providers [14-16], prompting them to seek out their own additional information or resources, often from internet communication technologies such as the internet and mobile apps.

The rapid global rise of internet communication technology provides many pregnant women with access to health information, including physical activity and exercise advice, outside of the traditional relationship with a health care provider [17-19]. For instance, in a cross-sectional survey of 293 pregnant US women, a little less than half (44%) had sought information on physical activity through the internet [20]. Similarly, the ubiquity of smartphone ownership in both high- and middle-income countries now allows most pregnant women to use mobile apps as a source of health information [18]. For example, a cross-sectional survey of 410 pregnant Australian women reported that almost three-quarters (73%) used at least one pregnancy app [21]. Although many women seek pregnancy-related information through web-based sources, few discuss this information with their health care providers [18,22].

Interestingly, there are more apps available for pregnancy than for any other medical topic [23]. This is of concern because health-related apps have previously been identified for their

XSL•FO
RenderX

potentially inaccurate content and poor quality [22,24,25]. A review of the quality of popular physical activity apps for the general population found that only 18% of the 65 included apps were based on consultation with an expert (eg, medical professional, fitness expert, or behavior change specialist) or on a peer-reviewed study [26].

Furthermore, despite it being well established that behavior change techniques (BCTs) are fundamental to supporting lifestyle change [27], a recent review found that apps designed to promote physical activity and exercise during pregnancy scarcely incorporated BCTs that have demonstrated efficacy for physical activity behavior change (eg, prompt review of behavioral goals) during pregnancy on their platform [27,28]. Another review of the quality and perceived impact of apps designed to address physical activity and exercise during and after pregnancy, which used the Mobile Application Rating Scale, reported that none of the 54 included apps specified goal setting, despite research showing goal setting to be one of the most effective BCTs among pregnant and postpartum women [29].

Although mobile apps are ideally placed to provide easily accessible information for pregnant women to support physical activity and exercise participation, there is also concern over the apps' safety and lack of regulation of content [18]. Furthermore, it is unclear whether commercial apps on physical activity and exercise during pregnancy align with current evidence-based recommendations [18,22,23].

Objective

Therefore, the aims of this study are to examine (1) alignment of the content with evidence-based recommendations for physical activity and exercise in pregnancy (ie, screening practices; physical activity and exercise prescription, including exercise frequency, intensity, time, and type (FITT principle); exercise considerations; and warning signs and symptoms to stop activity during pregnancy); (2) delivery, format, and features of physical activity and exercise instruction; and (3) credentials of the app developers.

Methods

Methodological Approach

The methodological approach used in this study was informed by previous app reviews [28,30,31] that explored app quality, features, and BCTs among apps designed to (1) improve diet, physical activity, and sedentary behavior in children and adolescents [31]; (2) provide nutritional advice to pregnant women [30]; and (3) promote prenatal physical activity and exercise [28].

Search Strategy

Systematic searches were conducted in the Australian App Store and Google Play Store in October 2020. Apps were identified using combinations of search terms relevant to pregnancy and exercise or physical activity (see Multimedia Appendix 1 for detailed search term combinations and strategy). Each search term combination was entered individually in the App Store and Google Play Store databases without any specified search categories, and search results were automatically ordered by the respective app store's *relevance* algorithm. That is, ordered by text relevance (ie, search term relevance to app title, keywords, and primary category) and user behaviors (ie, number of downloads and user ratings).

Inclusion Criteria and Selection Process

The apps underwent an initial screening and were included if the title and brief description of the app suggested a focus on physical activity or exercise during pregnancy, was available in English, not used as a studio-booking tool, and did not require any external devices (eg, Kegel device, activity monitor, or physical books). App characteristics, including app name, developer, version, store (App Store or Google Play Store), category, year of last update, cost, and average user rating were then extracted from the remaining apps (Table 1). The apps then underwent a secondary screening for inclusion by 2 independent reviewers (KA and SC), as per best practice for systematic reviews [32], and were deemed eligible for inclusion in this review if (1) they had been published or updated since 2018 (to ensure currency), (2) they were free to download and did not require an immediate paid subscription, and (3) they had an average user rating of ≥4 out of 5 because apps with higher standardized user ratings are more frequently downloaded [33]. Any disagreements in the screening process were resolved by consensus.

Each of the eligible apps was then independently reviewed by 2 of the 22 reviewers (ie, the authors, who are recognized as having expertise in physical activity and exercise and pregnancy as well as app reviews, including researchers, health professionals, and clinicians such as exercise physiologists). This review involved downloading the app, user testing, and assessing app features and quality criteria. If an app offered a free trial of a premium version, the reviewers were asked to assess the content delivered in the free trial. If no free trial was offered or if the app did not have a premium version, the standard (free) content was assessed. Freemium content (ie, extra content at a cost) was not assessed, and apps requiring immediate paid subscription (ie, no free trial) were excluded. The reviewers were provided with fictitious profiles to be used when personal information was required as well as instructions to gain familiarity with the app before data extraction (Multimedia Appendix 2). In cases of disagreement between the 2 reviewers, a third reviewer was assigned to provide an additional review, specifically focusing on the item of disagreement, to arrive at a majority decision.

XSL•FO
RenderX

Table 1. Characteristics of the apps included in this review (N=27).

Store and app ID	App name	Developer	Costing[a]	Version	Update	Rating, mean	Ratings, n
App Store							
1	Baby2Body: Pregnancy Wellness	Baby2Body Limited	Subscription	3.5.7	2020	4.6	534
2	Emily Skye FIT: Workout App	Loup Pty Ltd	Subscription	1.18.0	2020	4.1	63
3	Juna: Pregnancy Workouts	Juna Media LLC	Freemium	1.9.5	2020	4.2	5
4	Moms Into Fitness	Moms Into Fitness, Inc	Subscription	5.801.1	2020	5	1
5	Pregnancy +	Health & Parenting Ltd	Freemium	5.11	2020	4.8	7400
6	Prenatal Yoga | Down Dog	Yoga Buddhi Co	Subscription	5.2.2	2020	4.9	1100
7	Tips for Pregnant: Hello Belly	HelloBaby, Inc	Freemium	2.0.9	2020	4.1	212
8	Tone It Up: Workout & Fitness	Tone It Up, LLC	Subscription	2.4.6	2020	4.7	127
9	U Pilates: Workouts & Exercise	U Pilates Ltd	Freemium	3.19.2	2020	5	1
10	Yoggy: pregnancy yoga workouts	Millefeuille Agency	Freemium	2.7	2020	4.6	29
11	YogiBirth: Pregnancy Yoga App	YogiBirth Pty Ltd	Freemium	1.1.3	2020	4.9	196
Google Play Store							
12	9MonthsGuide	9MonthsGuide Team	Free	4	2019	4.4	511
13	Happy Pregnancy App	Dr Sachin Gothi (ObGyn)	Free	7	2019	4.7	211
14	Healthy pregnancy tips	My Apps Studio	Free	1	2019	5	180
15	I'm Pregnant: Pregnancy Week By Week	BabyJoyApp	Freemium	4	2018	4.7	36,000
16	Jillian Michaels: The Fitness App	EM Digital LLC	Subscription	3.9.9	2020	4.5	4000
17	Kegel Exercises for Men & Women: A How-to Guide	MasterpieceApps	Free	3.1	2020	4.2	22
18	Move Your Bump	Move Your Bump	Freemium	5.900.1	2020	4.7	14
19	My pregnancy calendar app: baby countdown timer	BabyInside	Freemium	2.0.9	2020	4.6	1000
20	My Pregnancy Journey	My Pregnancy Journey	Freemium	1.0.9	2019	4.3	20
21	pregnancy calendar	ruthie apps	Free	4.11.7.0	2018	4.2	37
22	Pregnancy Companion: Week by Week Tracking	Healthcare Apps	Free	1.8	2020	4.3	253
23	Pregnancy Exercise and Workout at Home	Pregnur Apps	Free	2.0.9	2020	4.6	3000
24	Pregnancy Exercises	B6Squad Dev.	Free	6	2020	4.1	123
25	Pregnancy Guide	ARVIRA DEV	Freemium	1.12	2019	4.6	7000
26	Pregnancy Guide App	EllStudiosApp	Free	3	2020	4.4	164
27	Pregnant. Pregnancy by week. Pregnancy calendar	rusakov77	Free	3.84	2019	4.4	190

[a]Subscription: free presubscription trial; freemium: offers premium content through in-app purchases; free: all content freely accessible.

Data Extraction

A tool was specifically created for data extraction purposes based on 2 of the most recently released evidence-based recommendations for physical activity and exercise during pregnancy [1,2]. Before the full data extraction process, 6 reviewers piloted the data extraction tool with 5 apps. Interrater consistency and feedback were considered during refinement of the extraction tool (see Multimedia Appendix 2 for final extraction tool detail).

All 22 reviewers used the final extraction tool to assess whether the app asked the user for any personal information about themselves (eg, age, height, and weight) or about their current pregnancy (eg, due date, current trimester, and singleton or multiple pregnancy). They also assessed the alignment of app content with evidenced-based recommendations, including disclaimers and terms and conditions; contraindication screening and information; exercise behavior, intention, and preferences; physical activity and exercise content (FITT principle of exercise); contraindications to physical activity and exercise

XSL•FO
RenderX

during pregnancy; safe pregnancy exercises or warning signs and symptoms to stop physical activity and exercise [1,2]; how the information was delivered (through still image, video, audio, text, etc); opportunities for users to modify exercises; and whether the app provided validation or references for its content. In addition to dichotomous responses (yes or no) and multiple-choice selections, the reviewers were provided with open-response textboxes to elaborate on their review or add further information. Any disagreements among the reviewers were resolved in consultation with a third reviewer.

Statistical Analyses

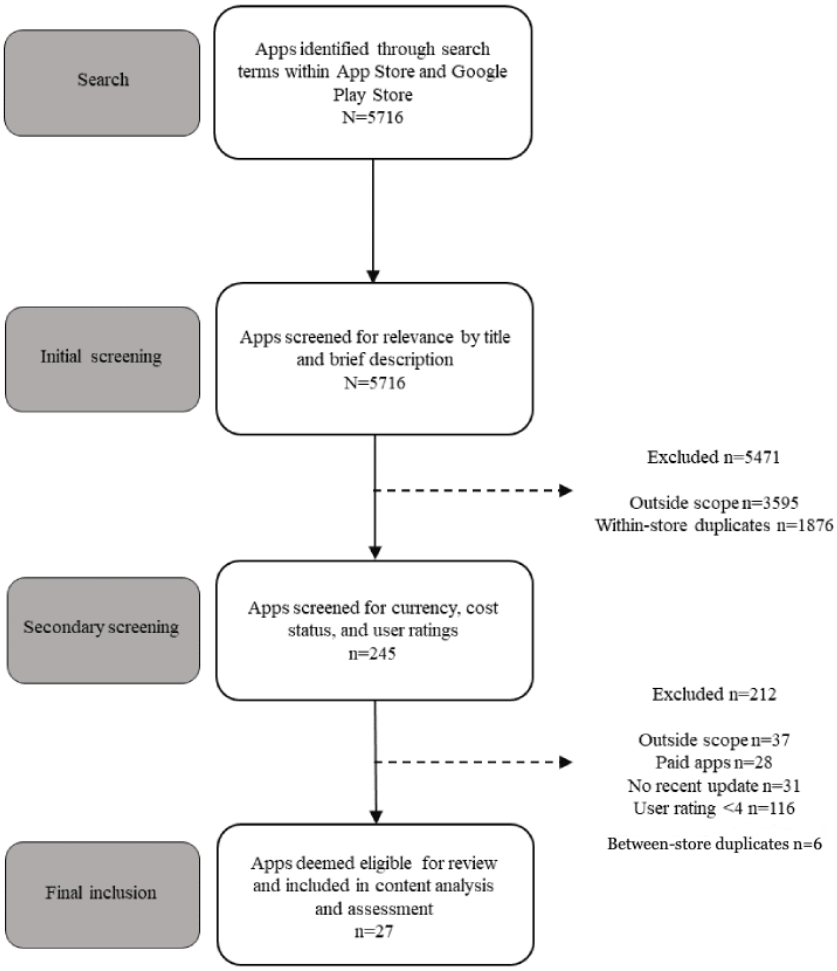
Descriptive statistics on the collated responses were derived (mean, SD, and frequency) using RStudio. Open-ended responses were summarized into *Other* categories for each representative item of interest.

Results

App Selection

A flowchart of the app selection process is presented in Figure 1. App Store and Google Play Store searches resulted in a total of 5716 apps for screening. The initial screening involved excluding 95.71% (5471/5716) of the apps that did not focus on physical activity or exercise in their title and description, were not available in English, required a studio-booking system, or required external devices. Of the remaining 245 apps, a second screening further excluded 212 (86.5%) apps based on the inclusion criteria and 6 (2.5%) apps that were between-store duplicates, leaving a total of 27 (11%) apps targeting exercise during pregnancy for inclusion in the final sample for data extraction and content analysis.

Figure 1. Flowchart of the app selection process.



App Characteristics

Of the 27 reviewed apps, 16 (59%) were accessed through Google Play Store and 11 (41%) through the App Store (Table 1). All apps were free to download; of the 27 apps, 10 (37%) offered all content for free, 6 (22%) offered a free trial of the premium subscription, and 11 (41%) offered restricted content in the absence of a paid premium version or paid subscription. The average star rating in the app stores at the time of downloading the apps was 4.54 of 5 (SD 0.29; range 4.1-5; skew 0.009; median 4.6, IQR 0.5), with a wide range of the number of users rating each app (range 1-36,000 user ratings; mean 2310.85, SD 7026.49; skew 4.09; median 190, IQR 734). All the included apps primarily prescribed structured exercise (ie, intentional and predetermined activity sessions), rather than lifestyle physical activity (ie, recreational activity incorporated into daily living, eg, gardening). Almost half of the apps were considered general pregnancy apps that provided a range of pregnancy information, including exercise-specific content; 33% (9/27) were considered pregnancy-specific exercise apps;

XSL•FO
RenderX

and 19% (5/27) were considered general exercise apps that offered a section specifically for pregnancy exercise content. Of the 27 apps, 18 (67%) presented users with several types of exercises (eg, a variety of aerobic, strength, pelvic floor, and flexibility exercises), whereas the remaining 9 (33%) offered users only 1 specific type of exercise (eg, only yoga or only weightlifting). Characteristics of the 27 apps included in this review are presented in Table 1.

Alignment of Content With Guidelines for Physical Activity and Exercise in Pregnancy

Disclaimers and Terms and Conditions

Of the 27 apps, 19 (70%) presented users with a disclaimer or terms and conditions absolving the app developers of liability (ie, user participation is at their own risk and the app is not responsible for any adverse outcomes that may occur when using the app or as a result of using the app). Users were required to actively agree to the terms and conditions in 37% (10/27) of the apps, whereas 19% (5/27) required users to agree to a disclaimer and 11% (3/27) asked users to actively agree to both terms and conditions as well as a disclaimer. Within the apps' disclaimer or terms and conditions, 67% (18/27) of the apps recommended that women seek medical clearance before commencing exercise during pregnancy (Multimedia Appendix 3). However, only 4% (1/27) specifically asked users to confirm whether they had obtained approval or clearance from their professional health care provider to engage in exercise while pregnant.

Screening for Contraindications to Exercise During Pregnancy

None of the apps specifically asked the user whether they had any absolute (eg, ruptured membranes, pre-eclampsia, or preterm labor) or relative (eg, symptomatic anemia or history of spontaneous miscarriage) contraindications to exercise during pregnancy.

Information about contraindications to exercise during pregnancy was limited. Of the 27 apps, 17 (63%) made no reference to contraindications, whereas the remaining 10 (37%) noted ≥1 recognized contraindications. The most frequently noted contraindications were history of spontaneous miscarriage, premature labor, or fetal growth restrictions (8/27, 30%), placenta previa (7/27, 26%), and persistent second- or third-trimester bleeding (7/27, 26%; Table 2 and Multimedia Appendix 4). Contraindications and related information were accessible not only within app information tabs and sections or within the exercise instructions and demonstrations, but also in some disclaimers and terms and conditions. Other medical issues of concern, such as chronic toxicities and infectious diseases, were noted in 15% (4/27) of the apps.

Table 2. Information about contraindications provided in the apps (N=27).

Contraindication	Apps providing information, n (%)
History of spontaneous miscarriage, premature labor, or fetal growth restriction	8 (30)
Placenta previa	7 (26)
Persistent second- or third-trimester bleeding	6 (22)
Serious cardiovascular, respiratory, or systemic disorder	5 (19)
Incompetent cervix	5 (19)
Ruptured membranes or premature labor	5 (19)
Pre-eclampsia	5 (19)
Mild to moderate cardiovascular or chronic respiratory disease	4 (15)
Symptomatic anemia	3 (11)
Evidence of intrauterine growth restriction	2 (7)
Multiple gestation (eg, triplets or higher number)	2 (7)
Pregnancy-induced hypertension	2 (7)
Poorly controlled type 1 diabetes, hypertension, or thyroid disease	1 (4)
Twin pregnancy after the 28th week	1 (4)

Exercise Behaviors, Intentions, and Preferences

Of the 27 reviewed apps, 4 (15%) asked the user questions about their current exercise behaviors and 4 (15%) asked the user questions about their current intentions for exercise (Multimedia Appendix 4). When closing the app and reopening it or logging back into the account, only 7% (2/27) of the apps asked the user whether they wanted to provide or update any personal information or exercise preferences (eg, weight, height, number of sessions per week, type of exercise session, and available equipment).

FITT Principle of Exercise

Of the 27 apps, only 5 (19%) recommended accumulating at least 150 minutes of exercise per week, whereas 6 (22%) specified the amount of exercise that should be accumulated per day. Of the 27 apps, 13 (48%) provided information on exercise frequency and 14 (52%) recommended exercise intensity, whereas only 4 (15%) provided information about the duration of each exercise session. All apps recommended at least one type of exercise (Table 3 and Multimedia Appendix 5).

XSL•FO
RenderX

Of the 27 apps, 13 (48%) recommended exercise frequency; all of them noted that exercise should be performed on most, if not all, days of the week in accordance with current evidence-based guidelines for physical activity and exercise during pregnancy. However, of these 13 apps, only 3 (23%) also suggested at least two sessions of resistance-based exercise per week.

Table 3. Recommended frequency, intensity, time, and types of exercises during pregnancy in the reviewed apps (N=27).

	Values, n (%)
Frequency of exercise	
Exercise on most, if not all, days of the week	13 (48)
Two sessions of resistance-based exercise per week	3 (11)
Intensity of exercise	
Light intensity	10 (37)
Moderate intensity	10 (37)
Vigorous intensity	4 (15)
Intensity measurement tool	
Talk Test to judge intensity of exercise	5 (19)
Borg Rating of Perceived Exertion Scale	4 (15)
Heart rate zones (based on age and fitness level)	3 (11)
Total time	
Accumulate at least 150 minutes of exercise per week	5 (19)
Exercise duration bouts (minutes)	
Exercise for 30 minutes per day	6 (22)
Exercise for at least 15 minutes per session	4 (15)
Break up exercise into small bouts	3 (11)
Type of exercise	
Aerobic	
Walking or jogging or running	11 (41)
Swimming	5 (19)
Cycling	3 (11)
Muscle strengthening	
Pelvic floor or Kegel exercises	19 (70)
Strength training (resistance or weight)	14 (52)
Other	
Yoga	21 (78)
Stretching or flexibility	19 (70)
Pilates	11 (41)
High-intensity interval training	5 (19)

Of the 27 apps, 14 (52%) recommended exercise intensity. Of these 14 apps, 10 (71%) recommended that users engage in light-intensity physical activity, whereas 10 (71%) recommended moderate-intensity physical activity (6/14, 43% suggested both light and moderate intensities). Specifically, of the 14 apps, only 4 (29%) recommended moderate- to vigorous-intensity physical activity in accordance with current guidelines [1,2]. Furthermore, of the 14 apps, only 6 (43%) suggested a measurement tool to monitor exercise intensity. These included the Talk Test (5/14, 36%) [34], the Borg Rating of Perceived Exertion Scale (4/14, 29%) [35], and heart rate zones (3/14, 21%) [36].

Of the 27 apps, only 9 (33%) recommended a total weekly duration of exercise, although guidelines for some physical activity and exercise during pregnancy suggest that women should work toward accumulating a total of 150-300 minutes of physical activity and exercise per week [1,2]. Of the 9 apps that did recommend a total duration, 5 (56%) recommended at least 150 minutes per week, whereas 6 (67%) recommended at least 30 minutes per day. Of these 9 apps, 3 (33%) also advised women that they could break up their exercise into smaller

XSL•FO
RenderX

bouts, if required, in accordance with current evidence-based guidelines for prenatal physical activity and exercise for pregnant women, which recommend that women progressively build their activity levels toward meeting the guidelines [1,2]. Conversely, of the 9 apps, 4 (44%) specifically suggested that exercise should be accumulated in bouts of at least 15 minutes each, which does not align with current guidelines for physical activity and exercise for pregnant women [1,2].

All apps recommended at least one type of exercise, with yoga (21/27, 78%), stretching or flexibility exercises (19/27, 70%), and pelvic floor or Kegel exercises (19/27, 70%) being the most frequently recommended. Aerobic exercises such as walking, jogging, and running (11/27, 41%); swimming (5/27, 19%); and cycling (3/27, 11%) were less frequently recommended. Muscle-strengthening exercises were recommended in 52% (14/27) of the apps, whereas high-intensity interval training (HIIT) was recommended in 19% (5/27) of the apps.

Of the 27 apps, only 3 (11%) provided accurate advice in full accordance with current evidence-based guidelines for physical activity and exercise during pregnancy, in relation to the FITT principle of exercise. Progression of exercise (Multimedia Appendix 5) was included in 37% (10/27) of the reviewed apps, with the most recommended progressions focused on steady progression toward physical activity and exercise guidelines (9/27, 33%) and modifications to exercise as the pregnancy progresses (8/27, 27%).

Exercises Considered Safe, Exercises to Avoid, and Other Exercise Considerations

Of the 27 apps, 19 (70%) listed exercises that are considered safe during pregnancy (Table 4 and Multimedia Appendix 6), such as pelvic floor exercises (15/19, 79%), aerobic exercises (13/19, 68%), and muscle-strengthening exercises (13/19, 68%). Pregnancy-specific classes were less frequently listed (8/19, 42%).

Table 4. Safe exercises, exercises to avoid, and warning signs or symptoms to stop exercise during pregnancy as provided in the apps (N=27).

	Values, n (%)
Exercises considered safe	
Pelvic floor exercises	15 (56)
Aerobic physical activity and exercise (walking, cycling, and swimming)	13 (48)
Muscle-strengthening exercises using body weight, weights, or resistance bands	12 (44)
Pregnancy-specific classes	8 (30)
Exercises considered unsafe	
Risk of falling (eg, exercise requiring balance, coordination, and agility)	12 (44)
Risk of contact or collision (eg, basketball and soccer)	9 (33)
Long periods of laying in the supine position	9 (33)
Heavy lifting (weights or lifting weight overhead)	6 (22)
Significant changes in pressure (eg, skydiving and scuba diving)	5 (19)
Exercise at high altitude	2 (7)
Long periods of standing still	1 (4)
Warning signs or symptoms to stop exercise	
Persistent dizziness or feeling faint, which does not resolve with rest	12 (44)
Persistent excessive shortness of breath, which does not resolve with rest	11 (41)
Regular painful uterine contractions	9 (33)
Vaginal bleeding	8 (30)
Heat stress or hyperthermia in first trimester	8 (30)
Chest pain	7 (26)
Persistent loss of fluid from the vagina (possible ruptured membrane)	7 (26)
Inadequate nutrition	7 (26)
Dehydration	6 (22)
Severe headache	5 (19)

Of the 27 apps, only 13 (48%) listed specific exercises to avoid or physical activities that are recognized as unsafe during pregnancy. These included physical activities that increase the risk of falling (12/13, 92%), physical activities with an increased risk of contact or collision (9/13, 69%), and exercise in the supine position (9/13, 69%). Other exercise activities such as strong stretches, twists and backbends, skiing, skating, and

XSL•FO
RenderX

bungee jumping were less frequently mentioned as physical activities to avoid.

Further evidence-based considerations relating to exercise during pregnancy were apparent in 67% (18/27) of the apps. The most common of these was the importance of staying well hydrated (15/27, 56%), whereas 33% (9/27) of the apps recommended that women wear appropriate clothing during exercise (Multimedia Appendix 6).

Warning Signs and Symptoms to Stop Activity During Pregnancy

Of the 27 apps, 18 (67%) listed warning signs and symptoms for stopping exercise during pregnancy in accordance with current evidence-based guidelines for physical activity and exercise during pregnancy (Table 4 and Multimedia Appendix 6). Persistent dizziness or feeling faint (12/18, 67%) and persistent excessive shortness of breath that does not resolve with rest (11/18, 61%) were the most common signs and symptoms listed in the reviewed apps.

Delivery, Format, and Features of Exercise Instruction

Of the 27 apps, 18 (67%) presented exercise as a series of individual exercises and 12 (44%) presented exercise as a series of workouts (with predetermined sequences of exercises; Multimedia Appendix 7). Of the 27 apps, 4 (15%) provided a combination of individual exercises and workouts, whereas 3 (11%) provided a combination of workouts and programs (predetermined sequences of workouts); only 1 (4%) app provided a 3-way combination of individual exercises, workouts, and programs. The most common format used by the apps to deliver exercise instruction was written cues (22/27, 81%), followed by video demonstrations (16/27, 59%), spoken cues (14/27, 52%), and finally, illustrations or still pictures to demonstrate the exercise (12/27, 44%). Of the 27 apps, 22 (81%) provided a combination of ≥2 of these instructional formats, 12 (44%) provided a combination of ≥3 instructional formats, and 3 (11%) provided a combination of all 4 instructional formats.

Of the 27 apps, 18 (67%) provided up-front details of the exercise session to help users decide which exercise to perform (Multimedia Appendix 7). These details most commonly included the duration of the exercise session (17/18, 94%), the type of exercise session (13/18, 72%), the suggested trimester in which to perform the exercise (11/18, 61%), and the equipment required to perform the exercise (10/18, 56%). Of the 18 apps, only 1 (6%) provided information up-front regarding the FITT principle of exercise, experience level, equipment, and trimester to fully inform the user of the exercise details. Moreover, only 17% (3/18) of these apps allowed users to modify the FITT parameters; however, none of these apps provided advice or feedback about exercise modifications.

Expertise and Credentials of App Developers

Of the 27 apps, only 6 (22%) specified the app developers' formal qualifications and 8 (30%) provided information to imply the developers' expertise or credibility (4/27, 15%, apps provided information on both formal and experiential credibility; Table 5 and Multimedia Appendix 7). Developer qualifications included master's in kinesiology, master's in exercise science, obstetrician, medical degree, master's in sports medicine, sports psychologist, qualified fitness instructor, personal trainer, master's in biology, Pilates instructor, midwife, and childbirth educator. The expertise or credibility of the app developer was often implied by the developer or instructor stating that they had years of practical experience in prenatal and postnatal support domains or had been through pregnancy themselves. Apps reporting formal qualifications or experiential credibility within their development team were more likely to support their content with recognized references. However, there was no difference in alignment of the content with evidence-based recommendations for physical activity and exercise in pregnancy.

Table 5. Developers' recognized credentials, experiential credibility, and referenced sources of information as provided in the apps (N=27).

	Apps providing developer information, n (%)
Developers' credentials and credibility	
Implies developers' credibility or experience	8 (30)
Specifies developers' recognized qualifications	6 (22)
Reference to recognized sources of information	
Government guidelines (exercise guidelines that incorporate pregnancy and guidelines for exercise during pregnancy)	7 (26)
Academic literature	6 (22)
Obstetrics-related guidelines	5 (19)

Sources Used to Guide App Content

Of the 27 apps, only 9 (33%) referenced a recognized or high-quality source of information (Table 5 and Multimedia Appendix 7), such as government (7/27, 26%) or obstetrics-oriented guidelines (5/27, 19%). In addition, 22% (6/27) of the apps referred to academic literature. However, of the 27 apps, only 1 (4%) provided a reference list to support the content provided.

Discussion

Principal Findings

This is the first review of apps that promote exercise in pregnancy to examine (1) alignment of the content with

XSL•FO
RenderX

evidence-based recommendations for physical activity and exercise in pregnancy; (2) delivery, format, and features of exercise instruction; and (3) credentials of the developers. Specifically, we identified a lack of alignment with current evidence-based recommendations for physical activity and exercise [1,2], particularly relating to screening of contraindications to exercise during pregnancy and to exercise prescription (FITT principle). In addition, few apps provided appropriate user opportunity to tailor the apps to their individual exercise needs, listed the sources of information used to guide content, or showed the credentials of the app developers. The results of this review highlight a need for improved regulation of the content of apps that promote exercise during pregnancy.

Current evidence-based guidelines for physical activity and exercise during pregnancy clearly indicate absolute and relative contraindications for participation in physical activity and exercise [1,2]. Women with absolute contraindications to physical activity and exercise during pregnancy are advised to avoid moderate to vigorous activity because the benefits of physical activity and exercise are outweighed by the risks [1]. Furthermore, those with relative contraindications are advised to discuss the advantages and disadvantages of physical activity and exercise, as well as potential modifications, with an appropriately qualified health care provider (such as their obstetric care provider) before participation [1]. Yet, almost two-thirds of the apps made no reference to absolute or relative contraindications and none included any kind of screening for pregnant women to identify contraindications. This is of concern, given that many pregnant women are increasingly turning to apps for guidance and support rather than relying on face-to-face information that they would traditionally receive from their health care providers [18]. Specifically, no users were asked to enter information or *check* contraindications from a provided list at any stage while engaging with the app. Although two-thirds of the apps advised users to seek medical clearance before commencing physical activity, none acknowledged (or considered) that contraindications to exercise can occur at any time throughout pregnancy. Thus, active screening for contraindications (using a simplified screening tool or method to limit impact when accessing the app) should be a feature of all apps and should be repeated with each user interaction frequently throughout the pregnancy. Warning signs and symptoms to stop physical activity and exercise are also clearly listed in current evidence-based guidelines for physical activity and exercise during pregnancy [1,2]. Despite this, only half of the apps provided educational information on these signs and symptoms. As such, women may continue to engage in physical activity and exercise while also risking the health and well-being of their pregnancy because they are unaware of the signs and symptoms to cease activity. The inclusion of a simple checklist listing the signs and symptoms to stop activity may help to prevent potential adverse events, while further providing evidence-based information to users. Furthermore, although 67% (18/27) of the apps included *Terms and Conditions* or a *Disclaimer* to encourage women to seek clearance from a health care provider before commencing exercise, only 44% (12/27) required acceptance of, or agreement with, the conditions or disclaimer. In fact, the American College of Obstetricians and Gynecologists recommends that a thorough clinical examination be conducted before commencing an exercise program to ensure the safety and well-being of the pregnancy and to check that pregnant women do not have any medical reasons to avoid exercise during this unique life stage [37].

Although most of the apps provided some level of information on frequency, intensity, total time (as well as duration of session bouts), or type of exercise, only 11% (3/27) of the apps did so in line with current guidelines. Instead, most of the reviewed apps recommended light-intensity exercise, although guidelines recommend that women engage in moderate- to vigorous-intensity physical activity and exercise [1,2] to achieve the greatest health benefits. In addition, 19% (5/27) of the apps recommended HIIT, which typically consists of alternating periods of vigorous- to high-intensity aerobic exercise with light recovery exercise or no exercise [38]. Although preliminary research suggests that this type of training seems to be well tolerated among a small cohort (N=14) of active pregnant women who engaged in a single session of HIIT [39], there is insufficient evidence to suggest that HIIT is safe during pregnancy. As such, current evidence-based guidelines do not recommend it [1,2]. Moreover, only 22% (6/27) of the apps provided women with advice on how to measure and monitor exercise intensity.

Very few apps provided users with up-front information pertaining to the exercise or workout (such as equipment required, duration and intensity of workout, and experience level), therefore limiting the users' ability to make an informed decision regarding the appropriateness of the exercise and or workout. By providing this important information up-front, the user can make an informed decision about the appropriateness of the exercise or workout, while taking into account their own exercise behaviors, experiences, and current pregnancy status.

Given that only one-third of the reviewed apps referred to relevant and recognized expert sources of information (ie, physical activity and exercise guidelines or peer-reviewed literature) within the app, it is not surprising that few aligned their exercise prescription with current evidence-based guidelines. These findings are similar to other app reviews that have evaluated the accuracy of app content [18,40]. For instance, Subhi et al [40] conducted a review of 52 studies (N=6520 apps) to examine expert involvement and adherence of app content to medical evidence in medical mobile phone apps. They found 30 studies (which included 3051 apps) that explored adherence to medical evidence in app content. In 17 of these studies, none of the app content was found to accurately reflect, or adhere to, medical evidence. The remaining 13 studies found that 10%-87% of the apps' content accurately reflected medical evidence. Moreover, only 5 of these 13 studies reported complete adherence and alignment of app content to medical evidence in more than 50% of the assessed apps [40].

Few apps collected users' individual activity characteristics, thus limiting the ability to appropriately tailor exercise prescriptions. Only half the included apps collected information about the current pregnancy, and only 15% (4/27) of the apps asked about current exercise behaviors. Given that these characteristics are fundamental to safe and appropriate exercise prescription in pregnancy [41], this may lead to women engaging

XSL•FO
RenderX

in inappropriate physical activity. Users' individual characteristics, including their current physical activity and exercise behaviors and medical history, are essential if apps are to support appropriate and individualized physical activity and exercise prescriptions that align with evidence-based guidelines [1,2].

Finally, only 22% (6/27) of the apps provided the formal qualifications of the app developers. This is consistent with previous reviews of the involvement of experts in app development [18,26]. For example, a recent review of popular physical activity apps for the general population found that only 12 of the 65 reviewed apps reported expert (ie, fitness expert, behavior change specialist, and medical professional) involvement in the development of the app [26]. Similarly, in the review by Subhi et al [40] of 52 studies, 28 studies assessed 3852 medical apps for expert involvement. The review found that 9%-67% of the apps reported expert involvement to some extent [40]. In a review of 129 urology apps to identify predictors of the number of urology app downloads, the explicit participation of urologists in app development was found to be likely to enhance the apps' chances to have a higher number of downloads [42], signifying the potential to improve app quality without compromising app popularity. However, it is important that those engaged in the app development process are up to date with current evidence-based guidelines because this review suggests that, despite involving those with formal qualifications or experiential credibility in the development process, the content was no more aligned with evidence-based recommendations for physical activity and exercise during pregnancy than in apps that did not report expert involvement.

Implications of Findings and Future Directions

It is clear that the rapid proliferation of apps targeting exercise in pregnancy has not been accompanied by a focus on ensuring user safety, adherence with evidence-based guidelines, and appropriateness of content. Given that health apps remain largely unregulated [18], there is a need for knowledge translation and implementation science to improve future practice. This should involve collaboration with stakeholders (ie, pregnant women) to ensure user satisfaction and with app developers, health care providers, and researchers to ensure that apps reflect evidence-based guidelines [43,44]. A key area of focus should be the incorporation of thorough pre-exercise screening practices that enable appropriate tailoring of exercise prescription to each user's unique individual characteristics. This may mean better incorporation of screening through the individual apps themselves or the creation of a stand-alone pre-exercise screening app to integrate with individual platforms. Apps that demonstrate collaboration or review by experts in physical activity and exercise during pregnancy could be recognized or registered through an app directory [43,44]. However, app directories to date do not include physical activity and exercise apps for pregnant women. Instead, they tend to focus on medical services such as streamlining communication among patients, providers, and their caregivers, allowing 24/7 management of a patient's condition and prescriptions, and improving organizational workflow. Moreover, none of the apps included in this review were developed by 1 of the 9 companies selected by the US Food and Drug Administration to participate in the development of the Software Pre-Cert Pilot Program [45]. Such regulation or certification of apps would help to ensure user safety [43,44]. However, it remains unclear whether commercial apps designed to target physical activity behaviors among pregnant women will require approval through this precertification program because they may not be recognized as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device" [45]. Thus, many of the apps, such as those included in this review, will likely continue to be developed in an unregulated commercial market.

Strengths and Limitations

The strengths of this review are that it included a systematic search for apps on both the App Store (Apple) and Google Play Store and that all reviewers were experts in physical activity and exercise in pregnancy. As per best practice for conducting systematic reviews [32], 2 independent reviewers (KA and SC) extracted data from each app using a pre-established and piloted extraction tool. Although app reviews are widely accepted for providing a snapshot in time, this approach may be considered a limitation, albeit unavoidable because apps require updating every 2 years. A decision was made to only include apps with freely available content in this review and those with a user rating of ≥4 out of 5. This was based on previous research that shows that few consumers are willing to pay for health apps and that consumers are more likely to download apps with a higher user rating [46]. However, the exclusion of numerous apps from this review may be considered a limitation because the findings cannot be generalized to apps that require immediate paid subscription or apps that provide freemium content or apps with user ratings of <4 out of 5. To improve inclusiveness, future studies might consider app exposure rates, download rates, and user comments as part of their inclusion criteria. In addition, if an app was available in both the App Store and Google Play Store, the App Store app was selected for inclusion in the study because fewer reviewers had access to Android devices. Although this is not necessarily a limitation, it should be acknowledged that some apps included in this study may also be accessible on the Google Play Store with different app characteristics.

Conclusions

Apps are a popular source of information and guidance for health behaviors during pregnancy, including physical activity and exercise. Our results demonstrate that neither do apps provide appropriate screening features to identify potential contraindications to exercise during pregnancy, nor do they provide content in accordance with current evidence-based physical activity guidelines or personalization. Overall, very few apps were found to have been developed by, or in conjunction with, experts (ie, health or medical professionals with expertise in prenatal exercise). This review emphasizes a critical need for development of evidence-based, tailored apps with greater regulation to minimize the potential risk of injury, complications, and adverse pregnancy outcomes for both mother and child.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 26 KB - mhealth_v10i1e31607_app1.docx]

Multimedia Appendix 2
Data extraction tool.
[DOCX File , 665 KB - mhealth_v10i1e31607_app2.docx]

Multimedia Appendix 3
Personal information, terms and conditions, and disclaimer.
[DOCX File , 29 KB - mhealth_v10i1e31607_app3.docx]

Multimedia Appendix 4
Contraindications and behavior and intention screening.
[DOCX File , 41 KB - mhealth_v10i1e31607_app4.docx]

Multimedia Appendix 5
Frequency, intensity, time (total time accumulated throughout the week), and type principle of exercise and progression.
[DOCX File , 43 KB - mhealth_v10i1e31607_app5.docx]

Multimedia Appendix 6
Safety, benefits, and considerations.
[DOCX File , 54 KB - mhealth_v10i1e31607_app6.docx]

Multimedia Appendix 7
Delivery, format, and features of exercise instruction.
[DOCX File , 34 KB - mhealth_v10i1e31607_app7.docx]

References

1. Mottola MF, Davenport MH, Ruchat S, Davies GA, Poitras VJ, Gray CE, et al. 2019 Canadian guideline for physical activity throughout pregnancy. Br J Sports Med 2018 Nov;52(21):1339-1346. [doi: 10.1136/bjsports-2018-100056] [Medline: 30337460]
2. Evidence-based physical activity guidelines for pregnant women. Australian Government Department of Health. 2020. URL: https://www.health.gov.au/resources/publications/evidence-based-physical-activity-guidelines-for-pregnant-women [accessed 2021-12-19]
3. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2021-12-19]
4. Atkinson L, Shaw RL, French DP. Is pregnancy a teachable moment for diet and physical activity behaviour change? An interpretative phenomenological analysis of the experiences of women during their first pregnancy. Br J Health Psychol 2016 Nov;21(4):842-858. [doi: 10.1111/bjhp.12200] [Medline: 27287546]
5. Phelan S, Phipps MG, Abrams B, Darroch F, Schaffner A, Wing RR. Practitioner advice and gestational weight gain. J Womens Health (Larchmt) 2011 Apr;20(4):585-591 [FREE Full text] [doi: 10.1089/jwh.2010.2316] [Medline: 21413898]
6. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013 Mar 03;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
7. Gjestland K, Bø K, Owe KM, Eberhard-Gran M. Do pregnant women follow exercise guidelines? Prevalence data among 3482 women, and prediction of low-back pain, pelvic girdle pain and depression. Br J Sports Med 2013 May;47(8):515-520. [doi: 10.1136/bjsports-2012-091344] [Medline: 22904295]
8. Mielke G, Crochemore-Silva L, Domingues MR, Silveira M, Bertoldi A, Brown W. Physical activity and sitting time from 16 to 24 weeks of pregnancy to 12, 24, and 48 months postpartum: findings from the 2015 Pelotas (Brazil) birth cohort study. J Phys Act Health 2021 Apr 15;18(5):587-593. [doi: 10.1123/jpah.2020-0351] [Medline: 33863852]

XSL•FO
RenderX

9. Hayman M, Short C, Reaburn P. An investigation into the exercise behaviours of regionally based Australian pregnant women. J Sci Med Sport 2016 Aug;19(8):664-668. [doi: 10.1016/j.jsams.2015.09.004] [Medline: 26481261]
10. Hesketh KR, Evenson KR. Prevalence of U.S. pregnant women meeting 2015 ACOG physical activity guidelines. Am J Prev Med 2016 Sep;51(3):e87-e89 [FREE Full text] [doi: 10.1016/j.amepre.2016.05.023] [Medline: 27544437]
11. Melton B, Marshall E, Bland H, Schmidt M, Guion W. American rural women's exercise self-efficacy and awareness of exercise benefits and safety during pregnancy. Nurs Health Sci 2013 Dec;15(4):468-473. [doi: 10.1111/nhs.12057] [Medline: 23663292]
12. Duncombe D, Wertheim EH, Skouteris H, Paxton SJ, Kelly L. Factors related to exercise over the course of pregnancy including women's beliefs about the safety of exercise during pregnancy. Midwifery 2009 Aug;25(4):430-438. [doi: 10.1016/j.midw.2007.03.002] [Medline: 18063253]
13. Stickford AS, Taylor EK, Rodriguez DV, Stroup S, Nunnery DL. Exercise behaviors and beliefs among pregnant women in rural communities. Am J Lifestyle Med 2021 Jul 02:155982762110265. [doi: 10.1177/15598276211026591]
14. Hayman M, Taranto M, Reaburn P, Alley S, Short CE. An investigation into regional medical practitioners' knowledge of exercise during pregnancy guidelines. Aust J Rural Health 2017 Dec;25(6):382-383. [doi: 10.1111/ajr.12364] [Medline: 28800191]
15. Hayman M, Reaburn P, Alley S, Cannon S, Short C. What exercise advice are women receiving from their healthcare practitioners during pregnancy? Women Birth 2020 Jul;33(4):e357-e362. [doi: 10.1016/j.wombi.2019.07.302] [Medline: 31466828]
16. Grenier L, Atkinson S, Mottola M, Wahoush O, Thabane L, Xie F, et al. Be healthy in pregnancy: exploring factors that impact pregnant women's nutrition and exercise behaviours. Matern Child Nutr 2021 Jan;17(1):e13068 [FREE Full text] [doi: 10.1111/mcn.13068] [Medline: 32705811]
17. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
18. Musgrave LM, Kizirian NV, Homer CS, Gordon A. Mobile phone apps in Australia for improving pregnancy outcomes: systematic search on app stores. JMIR Mhealth Uhealth 2020 Nov 16;8(11):e22340 [FREE Full text] [doi: 10.2196/22340] [Medline: 33196454]
19. Sayakhot P, Carolan-Olah M. Internet use by pregnant women seeking pregnancy-related information: a systematic review. BMC Pregnancy Childbirth 2016 Mar 28;16:65 [FREE Full text] [doi: 10.1186/s12884-016-0856-5] [Medline: 27021727]
20. Huberty J, Dinkel D, Beets MW, Coleman J. Describing the use of the internet for health, physical activity, and nutrition information in pregnant women. Matern Child Health J 2013 Oct;17(8):1363-1372. [doi: 10.1007/s10995-012-1160-2] [Medline: 23090284]
21. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]
22. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
23. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
24. BinDhim NF, Trevena L. There's an app for that: a guide for healthcare practitioners and researchers on smartphone technology. Online J Public Health Inform 2015 Jul 1;7(2):e218 [FREE Full text] [doi: 10.5210/ojphi.v7i2.5522] [Medline: 26392848]
25. BinDhim NF, Trevena L. Health-related smartphone apps: regulations, safety, privacy and quality. BMJ Innov 2015 Mar 05;1(2):43-45. [doi: 10.1136/bmjinnov-2014-000019]
26. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
27. Currie S, Sinclair M, Murphy MH, Madden E, Dunwoody L, Liddle D. Reducing the decline in physical activity during pregnancy: a systematic review of behaviour change interventions. PLoS One 2013 Jun 14;8(6):e66385 [FREE Full text] [doi: 10.1371/journal.pone.0066385] [Medline: 23799096]
28. Hayman M, Alfrey K, Cannon S, Alley S, Rebar AL, Williams S, et al. Quality, features, and presence of behavior change techniques in mobile apps designed to improve physical activity in pregnant women: systematic search and content analysis. JMIR Mhealth Uhealth 2021 Apr 07;9(4):e23649 [FREE Full text] [doi: 10.2196/23649] [Medline: 33825693]
29. Tinius R, Polston M, Bradshaw H, Ashley P, Greene A, Parker A. An assessment of mobile applications designed to address physical activity during pregnancy and postpartum. Int J Exerc Sci 2021 Apr 1;14(7):382-399 [FREE Full text] [Medline: 34055180]

XSL•FO
RenderX

30. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. Matern Child Nutr 2019 Jul;15(3):e12768 [FREE Full text] [doi: 10.1111/mcn.12768] [Medline: 30569549]
31. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. Int J Behav Nutr Phys Act 2017 Jun 24;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
32. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
33. Biviji R, Vest JR, Dixon BE, Cullen T, Harle CA. Factors related to user ratings and user downloads of mobile apps for maternal and infant health: cross-sectional study. JMIR Mhealth Uhealth 2020 Jan 24;8(1):e15663 [FREE Full text] [doi: 10.2196/15663] [Medline: 32012107]
34. Persinger R, Foster C, Gibson M, Fater D, Porcari J. Consistency of the talk test for exercise prescription. Med Sci Sports Exerc 2004 Sep;36(9):1632-1636. [Medline: 15354048]
35. Borg G. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
36. Mottola M, Davenport M, Brun C, Inglis S, Charlesworth S, Sopper M. VO2peak prediction and exercise prescription for pregnant women. Med Sci Sports Exerc 2006 Aug;38(8):1389-1395. [doi: 10.1249/01.mss.0000228940.09411.9c] [Medline: 16888450]
37. Physical activity and exercise during pregnancy and the postpartum period: ACOG committee opinion, number 804. Obstet Gynecol 2020 Apr;135(4):e178-e188. [doi: 10.1097/AOG.0000000000003772] [Medline: 32217980]
38. Gibala M. High-intensity interval training: a time-efficient strategy for health promotion? Curr Sports Med Rep 2007 Jul;6(4):211-213. [Medline: 17617995]
39. Anderson J, Pudwell J, McAuslan C, Barr L, Kehoe J, Davies GA. Acute fetal response to high-intensity interval training in the second and third trimesters of pregnancy. Appl Physiol Nutr Metab 2021 Dec;46(12):1552-1558 [FREE Full text] [doi: 10.1139/apnm-2020-1086] [Medline: 34433004]
40. Subhi Y, Bube SH, Rolskov Bojsen S, Skou Thomsen AS, Konge L. Expert involvement and adherence to medical evidence in medical mobile phone apps: a systematic review. JMIR Mhealth Uhealth 2015 Jul 27;3(3):e79 [FREE Full text] [doi: 10.2196/mhealth.4169] [Medline: 26215371]
41. Mottola M. Components of exercise prescription and pregnancy. Clin Obstet Gynecol 2016 Sep;59(3):552-558. [doi: 10.1097/GRF.0000000000000207] [Medline: 27135872]
42. Pereira-Azevedo N, Osório L, Cavadas V, Fraga A, Carrasquinho E, Cardoso de Oliveira E, et al. Expert involvement predicts mHealth app downloads: multivariate regression analysis of urology apps. JMIR Mhealth Uhealth 2016 Jul 15;4(3):e86 [FREE Full text] [doi: 10.2196/mhealth.5738] [Medline: 27421338]
43. Magrabi F, Habli I, Sujan M, Wong D, Thimbleby H, Baker M, et al. Why is it so difficult to govern mobile apps in healthcare? BMJ Health Care Inform 2019 Nov;26(1):e100006 [FREE Full text] [doi: 10.1136/bmjhci-2019-100006] [Medline: 31744843]
44. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. J Am Med Inform Assoc 2020 Feb 01;27(2):330-340 [FREE Full text] [doi: 10.1093/jamia/ocz175] [Medline: 31599936]
45. Digital health software precertification (pre-cert) program. U.S. Food & Drug Administration. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-program [accessed 2021-12-18]
46. Liew MS, Zhang J, See J, Ong YL. Usability challenges for health and wellness mobile apps: mixed-methods study among mHealth experts and consumers. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e12160 [FREE Full text] [doi: 10.2196/12160] [Medline: 30698528]

Abbreviations

BCT: behavior change technique
FITT: frequency, intensity, time, and type
HIIT: high-intensity interval training

XSL•FO
RenderX

Edited by L Buis; submitted 27.06.21; peer-reviewed by ASY Lee , HL Tam; comments to author 10.09.21; revised version received 01.10.21; accepted 17.11.21; published 19.01.22.

Please cite as:
Hayman MJ, Alfrey KL, Waters K, Cannon S, Mielke GI, Keating SE, Mena GP, Mottola MF, Evenson KR, Davenport MH, Barlow SA, Budzynski-Seymour E, Comardelle N, Dickey M, Harrison CL, Kebbe M, Moholdt T, Moran LJ, Nagpal TS, Schoeppe S, Alley S, Brown WJ, Williams S, Vincze L
Evaluating Evidence-Based Content, Features of Exercise Instruction, and Expert Involvement in Physical Activity Apps for Pregnant Women: Systematic Search and Content Analysis
JMIR Mhealth Uhealth 2022;10(1):e31607
URL: https://mhealth.jmir.org/2022/1/e31607
doi:10.2196/31607
PMID:35044318

©Melanie J Hayman, Kristie-Lee Alfrey, Kim Waters, Summer Cannon, Gregore I Mielke, Shelley E Keating, Gabriela P Mena, Michelle F Mottola, Kelly R Evenson, Margie H Davenport, S Ariel Barlow, Emily Budzynski-Seymour, Natalie Comardelle, Madison Dickey, Cheryce L Harrison, Maryam Kebbe, Trine Moholdt, Lisa J Moran, Taniya S Nagpal, Stephanie Schoeppe, Stephanie Alley, Wendy J Brown, Susan Williams, Lisa Vincze. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 19.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Promoting Physical Activity and Weight Loss With mHealth Interventions Among Workers: Systematic Review and Meta-analysis of Randomized Controlled Trials

Jiyeon Jung[1], PhD; Inhae Cho[2], PhD

[1]Department of Nursing, Korea National Open University, Seoul, Republic of Korea
[2]College of Nursing, Korea University, Seoul, Republic of Korea

Corresponding Author:
Inhae Cho, PhD
College of Nursing
Korea University
145 Anam-ro Sungbuk-gu
Seoul, 02841
Republic of Korea
Phone: 82 2 3290 4912
Fax: 82 2 928 9108
Email: inhae05@gmail.com

Abstract

Background: Physical activity (PA) is a vital factor in promoting health in the workforce. Mobile health (mHealth) interventions have recently emerged in workplace health promotion as an effective strategy for inducing changes in health behaviors among workers; however, the effectiveness of mHealth interventions in promoting PA and weight loss for workers is unclear.

Objective: This study aims to provide a comprehensive analysis of current evidence on the effectiveness of mHealth interventions in promoting PA and weight loss among workers.

Methods: We searched relevant databases, including PubMed, Embase, CINAHL Complete, and the Cochrane Library, for publications on mHealth interventions in the English or Korean language from inception to December 2020. Randomized controlled trials that evaluated the effectiveness of mHealth in improving PA and weight loss were retrieved. A meta-analysis with a random effects model and subgroup analyses was performed on PA types and mHealth intervention characteristics.

Results: A total of 8 studies were included in this analysis. More than half of the studies (5/8, 63%) were identified as having a high risk of bias. The mHealth intervention group showed a significant improvement in PA (standardized mean difference [SMD] 0.22, 95% CI 0.03-0.41; $P<.001$; $I^2=78\%$). No significant difference in weight loss was observed when comparing the intervention group with the control groups (SMD 0.02, 95% CI –0.07 to 0.10; $P=.48$; $I^2=0\%$). A subgroup analysis was also performed; walking activity (SMD 0.70, 95% CI 0.21-1.19; $P<.001$; $I^2=83.3\%$), a multicomponent program (SMD 0.19, 95% CI 0.05-0.33; $P=.03$; $I^2=57.4\%$), objective measurement (SMD 0.58, 95% CI 0.05-1.10; $P<.001$; $I^2=87.3\%$), and 2 or more delivery modes (SMD 0.44, 95% CI 0.01-0.87; $P<.001$; $I^2=85.1\%$) were significantly associated with an enhancement in PA.

Conclusions: This study suggests that mHealth interventions are effective for improving PA among workers. Future studies that assess long-term efficacy with a larger population are recommended.

(JMIR Mhealth Uhealth 2022;10(1):e30682) doi:10.2196/30682

KEYWORDS

mHealth; physical activity; obesity; weight loss; workforce; workplace health promotion; mobile phone

XSL•FO
RenderX

Introduction

Background

The promotion of physical activity (PA) is reported to be a key strategy for health promotion. Regular PA is proven to help prevent and treat noncommunicable diseases, such as cancer, cardiovascular disease, diabetes, stroke [1], and cardiovascular disease mortality [2]. It can also prevent hypertension [3] and obesity [4] and improve health-related quality of life [5].

According to the World Health Organization, 25% of adults do not currently meet the PA recommendations [6]. Thus, the World Health Organization provides a global action plan and framework for practical and feasible policy actions to support, maintain, and increase PA [6]. Establishing and maintaining healthy lifestyles in the adult population is essential [7], and it should be noted that most of the adult population are workers [8]. Inadequate PA is identified as a significant problem in adult worker groups [9,10]. This is mainly owing to the decrease in the amount of nonwork activity of blue-collar workers and white-collar workers who have sedentary behavior during work [11].

Most employed adults spend a large part of their waking hours at work [12]; thus, workplaces provide a unique and fruitful health promotion setting that can significantly increase PA and potentially influence workers' health [13]. In addition, promoting workers' PA was reported to be potentially beneficial, improving health status and psychological well-being and increasing economic benefits for employers through increased productivity [14,15]. However, there are several barriers to PA, of which one of the most widely mentioned is a lack of time [15].

Improving PA through mobile technology (mobile health [mHealth]) is emerging as a major trend in workplace health promotion for interventional change [16]. mHealth is based on wireless devices and sensors that people wear during their daily activities, including mobile phones and is reported to be convenient and effective in changing health behavior [17,18]. In particular, it is recognized as a tool for intervention delivery that enables continuous monitoring during daily life and various interventions [10], thus enhancing one's responsibility for their own health and performance [19]. The proper use of mobile technologies for promoting PA may be a cost-effective and feasible way to reach this population [20].

Previous studies have investigated the use of mHealth to promote PA in various populations, including workers [21-31]. A study on mHealth apps and self-determination theory showed increased PA levels in motivated workers [30]. In addition, a large population-based mHealth intervention study reported significant improvements in PA, sitting times, and body weight [31]. mHealth devices not only track data but also encourage workers to achieve their health goals through sustained engagement [32]. A previous review concluded that mHealth interventions are potentially effective and feasible for increasing PA in the workplace [33], with some evidence of short-term weight loss [34]. In contrast, other studies reported nonsignificant changes in PA level [35] and weight control [7,24] in certain groups of workers.

There is a knowledge gap on the effectiveness of mHealth technologies in promoting PA [30] and weight loss [24] among workers. Furthermore, findings from the current literature are still inconclusive [33]. There is still some debate about the effectiveness of mHealth interventions in the working population.

Objective

In this study, we aim to provide a comprehensive analysis of current evidence from randomized controlled trials (RCTs) on the effectiveness of mHealth interventions in promoting PA and weight loss among workers.

Methods

Study Design

This study is a systematic review and meta-analysis of RCTs conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [36].

Search Strategy

A literature review of 4 bibliographic electronic databases—PubMed, Embase, Cochrane Library, and CINAHL Complete—was conducted. Published articles on mHealth from its inception until December 2020 were identified. We confirmed the search terms based on our research question. According to search terms, the Medical Subject Headings terms; Emtree, the related entry term; and free terms were collected from relevant articles and bibliographic databases. The keywords identified were as follows: *telemedicine*, *cell phone*, *smartphone*, *mobile device*, *mHealth*, *mobile applications*, *mHealth program*, *worker*, *employee*, *occupation worksite*, *working adult*, *workplace*, *occupational health*, *randomized controlled trial*, *clinical trial*, *controlled clinical trial*, *evaluation study*, and *quasi-experimental*. Our search strategies are presented in Multimedia Appendix 1.

After the search, relevant identified articles were exported using the bibliography software Endnote (Version X9.1; Clarivate Analytics) and duplicate papers were removed. The titles and abstracts were screened by 2 reviewers (JJ and IC) independently, using preset criteria; irrelevant publications were excluded and full-text articles were then selected. To identify additional studies, we manually checked the reference lists of relevant reviews found in the original search. The entire process, from developing a search strategy to selecting studies and cross-checking all publications, was carried out by the 2 reviewers (JJ and IC). In cases of inconsistent selection, an agreement was reached through discussion.

Inclusion and Exclusion Criteria

This study's eligibility criteria were specified according to the purpose of this review. On the basis of the participants, intervention, comparison, outcome, and study design framework, the inclusion criteria were as follows: (1) participants, working population and those aged ≥18 years; (2) intervention, any mHealth intervention that promoted PA using mobile

XSL•FO
RenderX

technologies (mHealth interventions were programs that used mobile phones with mobile functions, such as phone call, message service, app, GPS, Bluetooth technology, and others); (3) comparison, control group should refer to participants who did not receive any intervention using mobile phones; (4) outcome, the study's outcome included PA (eg, self-reported or device-reported PA, walking time, and the number of steps) or body weight to verify the effects of mHealth interventions on promoting PA and weight reduction in workers; and (5) study design, only the RCT design was considered.

The study's exclusion criteria included the following: (1) studies not published in English or Korean, (2) studies that targeted participants with a disease, (3) studies that reported incomplete or insufficient data (eg, study protocols, ongoing studies, and conference abstract), and (4) studies with web-based mHealth intervention.

Data Extraction

The data from the eligible studies were extracted using Excel (Microsoft Corporation) and were coded using a predesigned template by the research team. The data included general study characteristics (eg, author, published year, country, setting, design, participants, their age, and comparator), intervention characteristics (eg, mHealth intervention delivery mode, category, intervention contents, behavior change techniques, and duration), and the study's result (eg, outcome variables).

It has been reported that mHealth interventions are often performed together with various intervention components in workers' health promotion programs [33]. Thus, we classified intervention into 2 different categories. The included studies were classified into a stand-alone mHealth intervention using mobile device only or a multicomponent intervention where the use of mHealth device was one of several intervention components in the programs (eg, face-to-face counseling, printed materials, offline education, and organizational support). Finally, for the coding of behavior change techniques, we used the Coventry, Aberdeen, and London-Refined taxonomy by Michie et al [37]. This 40-item taxonomy can be used to systematically classify PA and healthy eating behaviors.

Data extraction was performed independently by the first reviewer (JJ) and confirmed by the second reviewer (IC). When discrepancies emerged, we resolved them through discussion until an agreement was reached.

Risk of Bias Assessment

The eligible studies were evaluated for the risk of bias using a revised Cochrane risk of bias tool. This tool was developed to assess the risk of bias in randomized trials [38]. The Cochrane risk of bias tool consists of the following five domains: (1) randomization process, (2) deviations from intended interventions, (3) missing outcome data, (4) measurement of the outcome, and (5) selection of the results. The risk of bias was evaluated using algorithms that depend on the answers to the questions in each domain. As a result, each domain was assigned 1 of 3 levels (*high risk*, *low risk*, and *some concerns*). The risk of bias in the included studies was assessed by combining the results across the domain responses. The 2 reviewers (JJ and IC) independently assessed the risk of bias in each article. If there were differences in evaluation between the 2 reviewers, they were discussed and resolved.

Statistical Analysis

The extracted data from the included studies were analyzed using Stata 17.0 (StataCorp LLC). We used a random effects model in this analysis. A meta-analysis was performed using continuous data. The standardized mean difference (SMD) was calculated as the Hedges *g* using mean and SDs. For extracted data without mean and SD, the Hedges *g* was estimated using other statistical data (eg, mean difference [MD], *P* value, and CI) using Comprehensive Meta-Analysis version 3 (Biostat Inc). The heterogeneity within selective studies was estimated using the statistic I^2 [39]. Subgroup analysis was conducted according to PA features, intervention category, the PA measurement, and the number of delivery modes of the mHealth program.

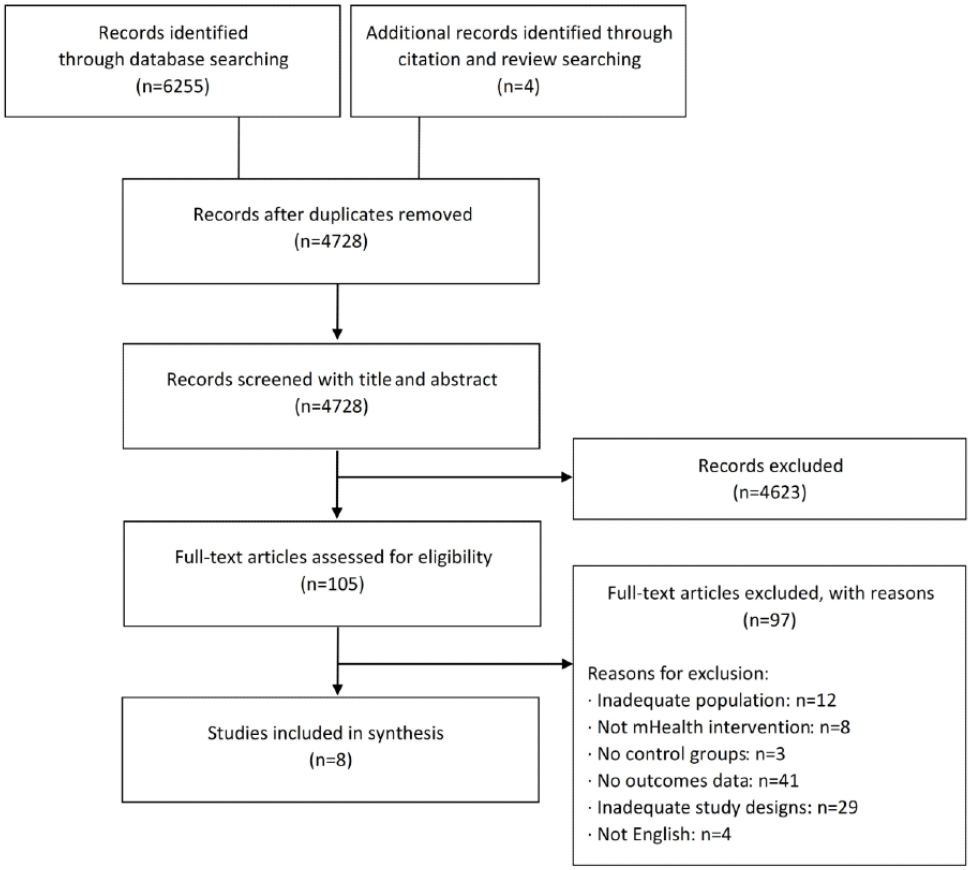
Results

Search Results

The search identified 6255 records in the bibliographic databases, and an additional 4 records were added through manual search from relevant reviews. After excluding duplicate records, the study titles and abstracts were screened; of the 6259 studies, 4623 (73.86%) studies that did not meet the eligibility criteria were excluded and the remaining 105 (1.67%) studies were checked. After a full-text review, 0.13% (8/6259) of the studies met the study eligibility criteria and were included in the meta-analysis (Figure 1) [22-29].

XSL•FO
RenderX

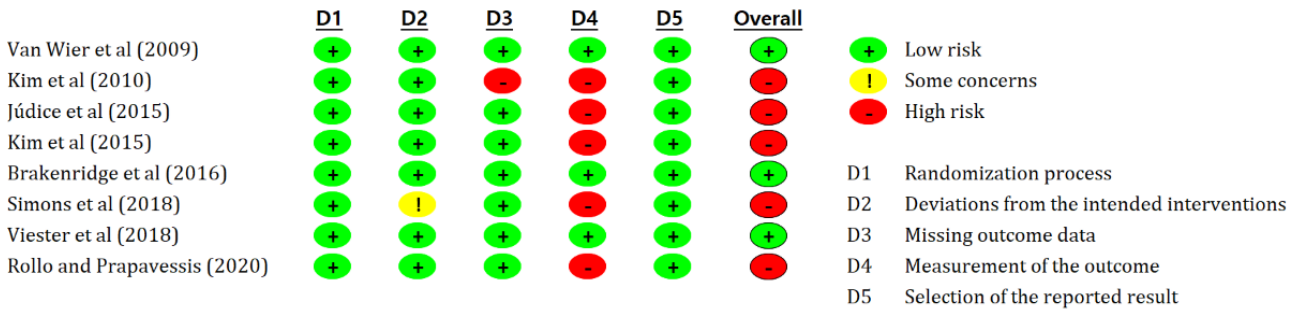
Figure 1. Flow diagram of study selection. mHealth: mobile health.



Risk of Bias

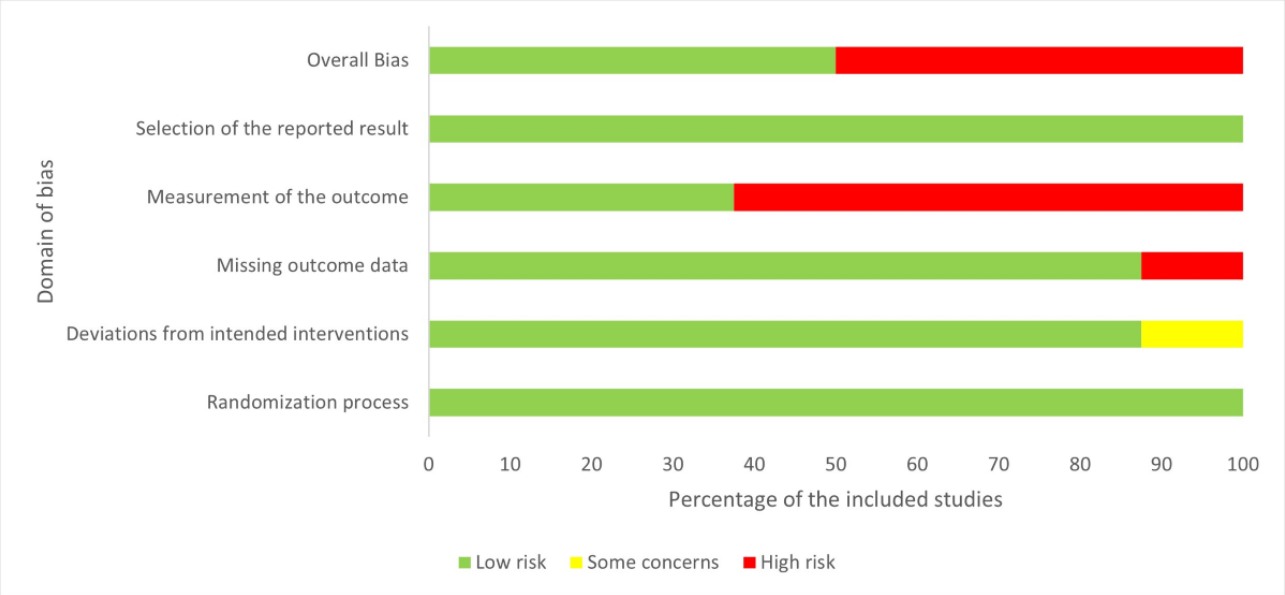
More than half of the studies (5/8, 63%) had a high risk of bias [23-25,27,29]. The risk of bias in the measurement of outcome was considered high in 63% (5/8) of the studies owing to self-report methods without sufficient blinding [23-25,27,29]. In the study by Kim et al [25], there was a high risk of bias for missing outcome data. All 8 studies had a low risk of bias in the randomization process and in selecting the reported result (Figures 2 and 3).

Figure 2. Results of risk of bias assessment for the included studies using Cochrane risk of bias tool 2.0 (detailed assessment of included studies) [22-29].



XSL•FO
RenderX

Figure 3. Results of risk of bias assessment for included studies using Cochrane risk of bias tool 2.0.



Study Characteristics

These meta-analysis studies were published between 2009 and 2020. Among the 8 studies, there were 2 (25%) studies each from the Netherlands and the United States and 1 (13%) study each from Canada, Portugal, Korea, and Belgium. Most studies had a 2-arm RCT design (5/8, 63%); the rest had a 3-arm RCT (1/8, 13%), cluster RCT (1/8, 13%), and crossover RCT (1/8, 13%). Most participants were healthy workers; only 13% (1/8) of the studies targeted obese employees (Table 1). Half of the included studies (4/8, 50%) focused only on PA, and the other studies (4/8, 50%) focused on the contents of PA and dietary change (Multimedia Appendix 2) [22-29].

XSL•FO
RenderX

Table 1. Characteristics of the included studies (N=8).

Study	Country	Setting	Design	Participants	Age (years), mean (SD)	Comparator	Outcomes	
							PA[a] (measurement)	BW[b] or BMI
van Wier et al [28]	Netherlands	7 companies (IT[c], hospitals, insurance, bank, and police force)	3-arm RCT[d]; IG[e]1: phone group; IG2: email group; CG[f]: control group	1386[g]; 929 (67.02%) male participants, 457 (32.97%) female participants; IG1 (n=462, 33.33%); IG2 (n=464, 33.47%); CG (n=460, 33.19%)	IG1: 43 (8.8); IG2: 43 (8.4); CG: 43 (8.7)	Printed materials	MET[h] minutes per week (SR[i]: SQUASH[j])	BW
Kim et al [25]	United States	43 companies and 13 community organizations	2-arm RCT	2470; IG (n=1279, 51.78%): 238 (18.6%) male participants, 1041 (81.39%) female participants; CG (n=1191, 48.22%): 246 (20.73%) male participants, 945 (79.34%) female participants	IG: 43.5 (10.3); CG: 43.6 (10.1)	Printed materials	Time of mild, moderate, and vigorous PA (SR: minutes per day)	BW
Júdice et al [23]	Portugal	Academic and administrative sectors of the university and others	Crossover RCT	10; 5 (50%) male participants, 5 (50%) female participants; IG (n=5, 50%); CG (n=5, 50%)	50.4 (11.5)	Usual care	Time of sitting and standing, walking time (hours per day), number of steps, and sitting time (OB[k]: ActivPAL)	N/A[l]
Kim et al [24]	Korea	3 public institutions	2-arm RCT	205 (n=196, 95.6% for analysis); 100% (196/196) male participants; IG (n=101, 51.5%); CG (n=95, 48.5%)	IG: 41.02 (6.82); CG: 41.55 (6.98)	Printed materials and face-to-face counseling	MET minutes per week (SR: IPAQ[m])	BW
Brakenridge et al [22]	United States	An international property and infrastructure company	Cluster RCT	153; 83 (54.2%) male participants, 70 (45.8%) female participants; IG (n=66, 43.1%); CG (n=87, 56.9%)	IG: 37.6 (7.8); CG: 40 (8)	Organizational support	Time of sitting and standing, walking time, and number of steps (OB: ActivPAL)	N/A
Simons et al [27]	Belgium	29 workplaces (shops, retail stores, catering industry, social employment businesses, factories, etc)	2-arm RCT	130; 63 (48.5%) male participants, 67 (51.5%) female participants; IG (n=60, 46.2%); CG (n=70, 53.8%)	IG: 24.8 (3.1); CG: 25.1 (3)	Printed materials	Time of light, moderate, and vigorous PA; MVPA[n] and total PA (OB: GT3X and accelerometers); occupational, household, recreational, active transport, and total PA (SR: IPAQ), and number of steps (OB: Fitbit Charge)	BMI
Viester et al [29]	Netherlands	A construction company	2-arm RCT	314; 100% (314/314) male participants; IG (n=162, 51.6%); CG (n=152, 48.4%)	IG: 46.3 (9.9); CG: 47 (9.5)	Usual care	Time of leisure-time MVPA (SR: SQUASH)	BW

XSL•FO
RenderX

Study	Country	Setting	Design	Participants	Age (years), mean (SD)	Comparator	Outcomes	
							PA[a] (measurement)	BW[b] or BMI
Rollo and Prapavessis [26]	Canada	Large businesses, office spaces, and universities	2-arm RCT	60; 5 (8%) male participants, 55 (92%) female participants; IG (n=29, 48%); CG (n=31, 52%)	IG: 46.59 (11.13); CG: 43.87 (11.54)	Usual care	Time of sitting and standing, walking time, and stretching (SR: OSPAQ[o])	N/A

[a]PA: physical activity.
[b]BW: body weight.
[c]IT: information technology.
[d]RCT: randomized controlled trial.
[e]IG: intervention group.
[f]CG: control group.
[g]Included overweight employees.
[h]MET: metabolic equivalent task.
[i]SR: self-reported.
[j]SQUASH: Short Questionnaire to Assess Health-Enhancing Physical Activity.
[k]OB: objective.
[l]N/A: not applicable.
[m]IPAQ: International Physical Activity Questionnaire.
[n]MVPA: moderate to vigorous physical activity.
[o]OSPAQ: Occupational Sitting and Physical Activity Questionnaire.

The mHealth Intervention

We identified the program characteristics of the included studies to confirm the features of mHealth interventions on PA among workers (Table 2). We classified them as follows: mHealth intervention delivery mode, intervention category, mHealth intervention contents, behavior change techniques, PA features, and duration (with or without follow-up).

The mHealth intervention delivery mode included phone calls, SMS text messages, wearable activity monitors, and smartphone apps. Half of the studies (4/8, 50%) included phone calls to motivate the participant to be physically active [22-24,28]; then, they were in the order of wearable activity monitors (3/8, 38%) [24,26,27], SMS text messages (3/8, 38%) [24,25,29], and apps (3/8, 38%) [25-27]. Half of the interventions (4/8, 50%) were implemented using 2 or more modes of delivery [24-27]. The intervention category was classified into multicomponent (6/8, 75%) [22,23,25,26,28,29] and stand-alone (2/8, 25%) [24,27]. The most used component in the multicomponent intervention was educational materials (5/8, 63%) [22,25,26,28,29]. The PA features were categorized into overall PA and walking activity. Of the 8 included studies, 4 (50%) studies reported overall PA [24,25,28,29], 3 (38%) studies reported walking activity [22,23,26], and 1 (13%) study dealt with both measurements [27]. The intervention duration ranged from 1 week to 12 months. Most studies spanned 12 months; only 3 (38%) studies reported a follow-up.

XSL•FO
RenderX

Table 2. Characteristics of mobile health interventions in the included studies (N=8).

Study	Delivery mode	Category	Mobile health intervention contents	Behavior change techniques	PA[a] features	Duration (follow-up)
van Wier et al [28]	Phone call	MC[b]: face-to-face counseling and educational materials	Phone call counseling, face-to-face counseling, and printed materials	Prompt self-monitoring, provide feedback, provide instruction, teach to use prompts, goal-setting, and provide information (printed materials)	Overall PA	6 months (no follow-up)
Kim et al [25]	Phone call	MC: educational materials	Phone call counseling and printed materials	Provide information (printed materials), goal-setting, action planning, problem solving, set graded tasks, prompt review, provide feedback, provide instruction, and stress management	Overall PA	6 months (no follow-up)
Júdice et al [23]	Phone call, SMS text message, and wearable activity monitor	SA[c]	Activity monitor, alert, and feedback	Goal-setting, prompt self-monitoring, teach to use prompts, and provide feedback	Walking activity	1 week (no follow-up)
Kim et al [24]	SMS text message	MC: offline education and face-to-face counseling	Tailored SMS text message, offline education, and face-to-face counseling	Goal-setting, problem solving, prompt self-monitoring, provide feedback, provide information, provide instruction, provide information (printed materials and face-to-face counseling), and use of follow-up prompts	Overall PA	6 months (no follow-up)
Brakenridge et al [22]	Wearable activity monitor and app	MC: organizational support (emails and educational materials)	Activity monitor, feedback, and organizational support	Prompt self-monitoring, provide feedback, plan social support, and provide information (printed materials)	Walking activity	12 months (no follow-up)
Simons et al [27]	Wearable activity monitor and app	SA	Activity monitor and feedback	Goal-setting, action planning, problem solving, set graded tasks, prompt review of behavioral goals, provide information, provide feedback, and prompt self-monitoring	Overall PA and walking activity	9 weeks (12 weeks)
Viester et al [29]	Phone call	MC: educational materials and organizational support	Phone call counseling, printed materials, and organizational support	Goal-setting, problem solving, prompt review of behavioral goals, provide information, provide feedback, prompt self-monitoring, and plan social support	Overall PA	6 months (12 months)
Rollo and Prapavessis [26]	SMS text message	MC: face-to-face counseling and educational materials	Tailored SMS text message, face-to-face counseling, and printed materials	Counseling, goal-setting, action planning, problem solving, set graded tasks, provide information, and teach to use prompts	Walking activity	6 weeks (8 weeks)

[a]PA: physical activity.

[b]MC: multicomponent.

[c]SA: stand-alone.

The Effects on PA and Weight Loss

All the 8 studies reported on PA [22-29]. Júdice et al [23], Brakenridge et al [22], and Simons et al [27] used more than 2 measurements as outcome variables; the results were included in the meta-analysis. These results were treated individually in the meta-analysis; therefore, 12 effects were analyzed in this PA meta-analysis. Results showed that the mHealth intervention group was significantly associated with an improvement in PA after completing the intervention compared with the control group (SMD 0.22, 95% CI 0.03-0.41; P<.001; I^2=78%).

Regarding weight loss in workers, 50% (4/8) of the studies, except the study by Simons et al [27] that did not report the results of body weight, were analyzed in the meta-analysis. There was no statistically significant difference in weight loss compared with control groups (SMD 0.02, 95% CI –0.07 to 0.10; P=.48; I^2=0%). A summary of the detailed findings is presented in Figures 4 and 5.

XSL•FO
RenderX

Figure 4. Meta-analysis of mobile health intervention effect on physical activity [22-29]. ES: effect size; OB: objective; PA: physical activity; SR: self-reported; WA: walking activity.

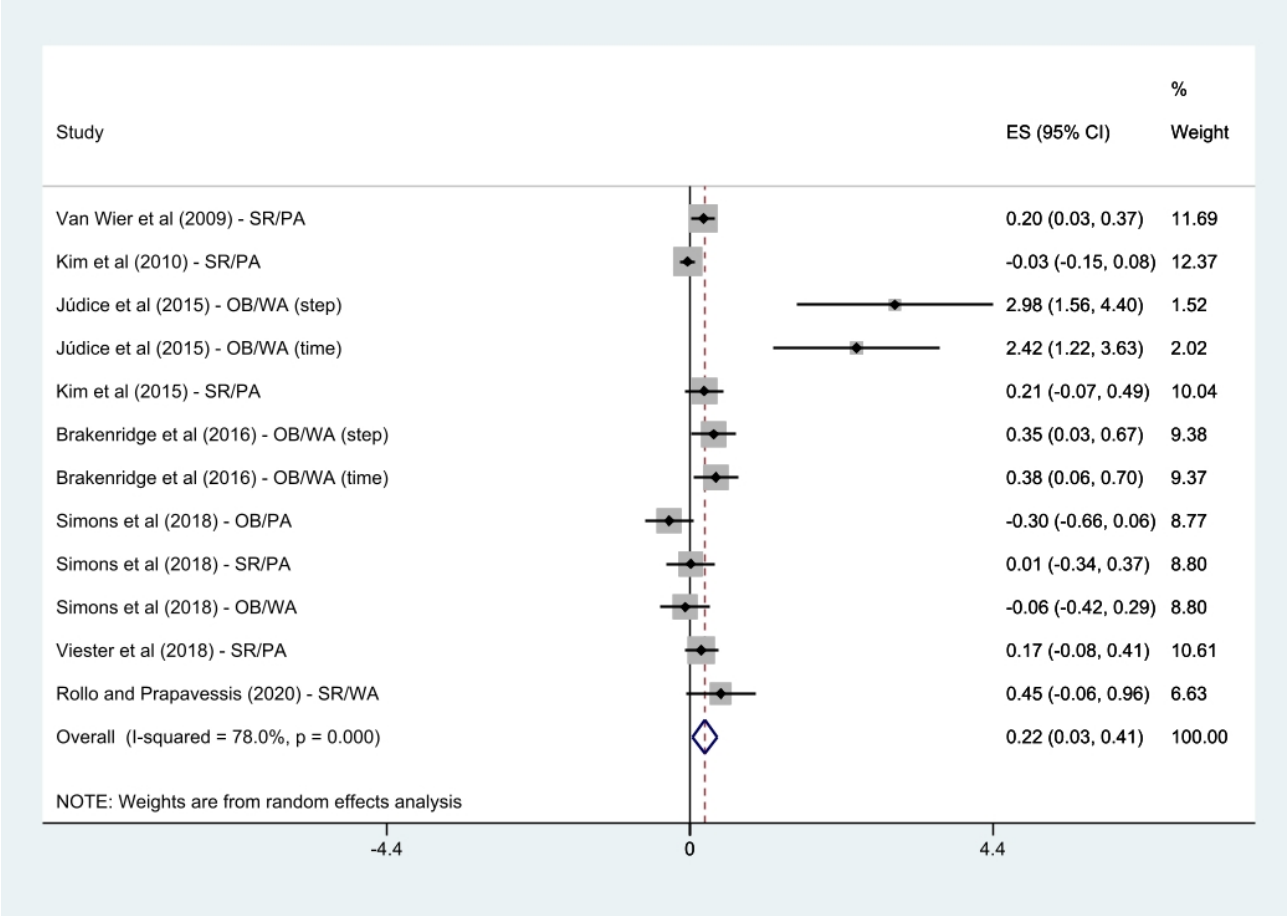
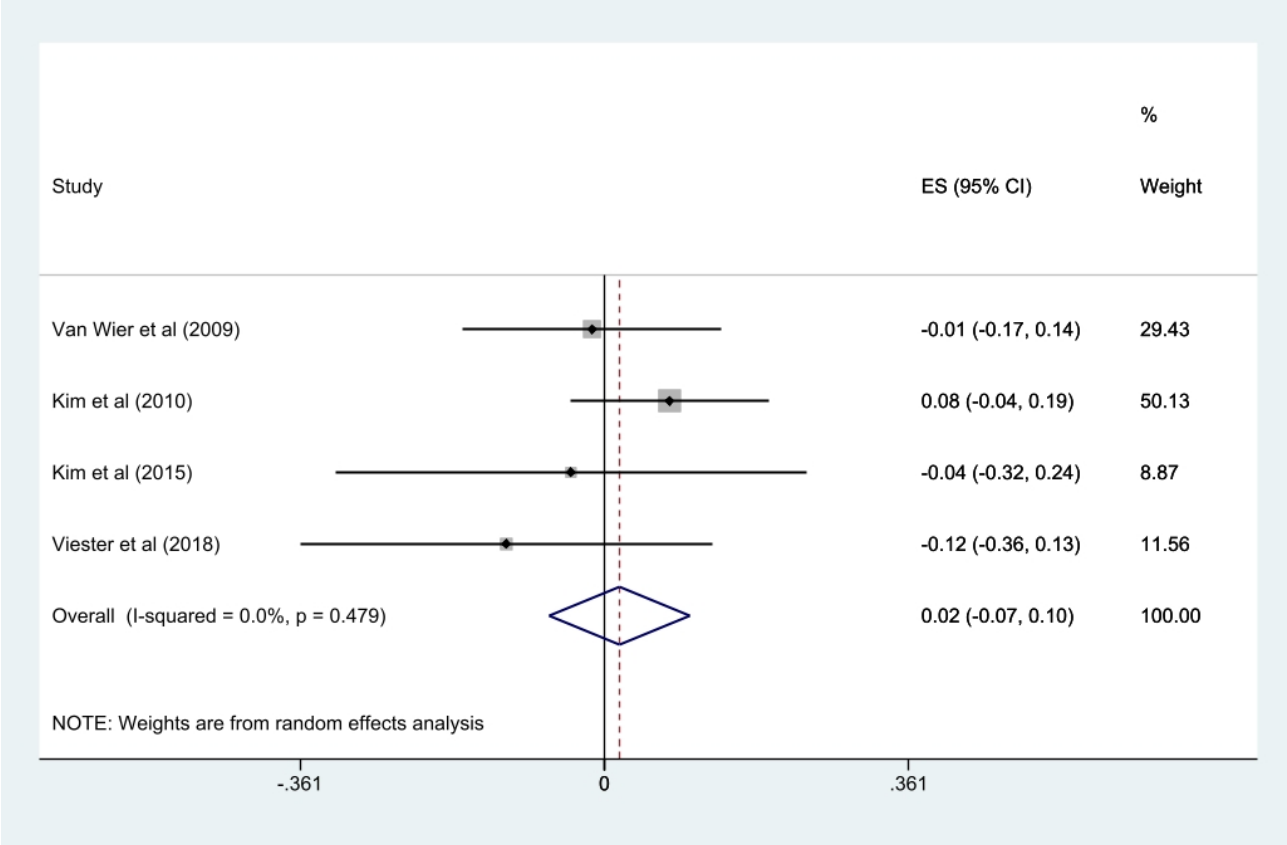


Figure 5. Meta-analysis of mobile health intervention effect on weight loss [24,25,28,29]. ES: effect size.



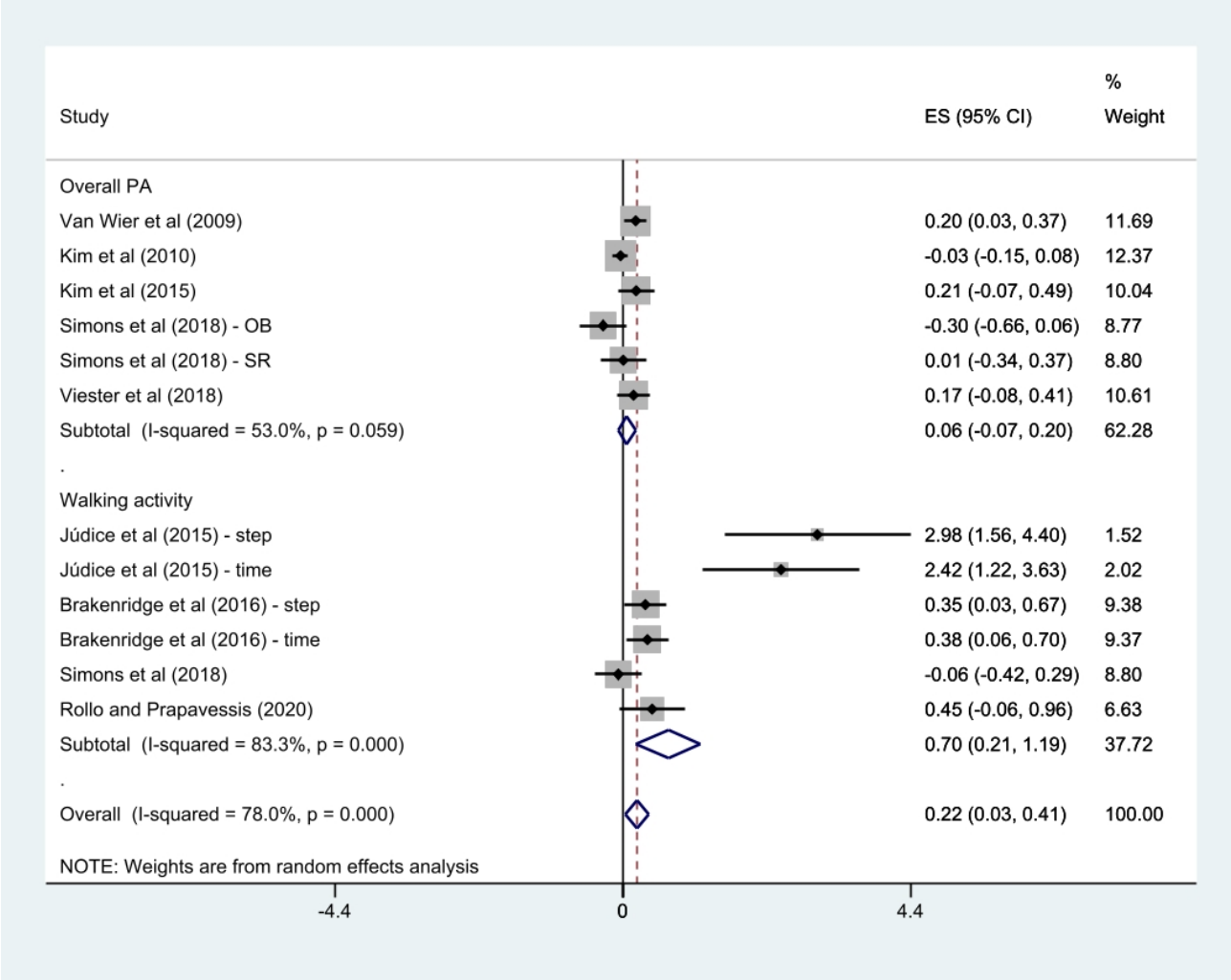
XSL•FO
RenderX

Subgroup Analysis

A subgroup analysis was conducted according to PA features (overall PA or walking activity), intervention categories (multicomponent program or stand-alone mHealth program), PA measurements (self-reported measurement or objective measurement), and the number of delivery modes (1, 2, or more).

The subgroups of walking activity (SMD 0.70, 95% CI 0.21-1.19; P<.001; I^2=83.3%), multicomponent program (SMD 0.19, 95% CI 0.05-0.33; P=.03; I^2=57.4%), objective measurement (SMD 0.58, 95% CI 0.05-1.10; P<.001; I^2=87.3%), and 2 or more delivery modes (SMD 0.44, 95% CI 0.01-0.87; P<.001; I^2=85.1%) showed a significant association with an enhancement in PA when compared with the control group. However, the overall PA (SMD 0.06, 95% CI –0.07 to 0.20; P=.06; I^2=53%), stand-alone mHealth program (SMD 0.63, 95% CI –0.05 to 1.32; P<.001; I^2=88.6%), self-reported measurement (SMD 0.12, 95% CI –0.01 to 0.25; P=.11; I^2=44.3%), and 1 delivery mode (SMD 0.14, 95% CI –0.01 to 0.28; P=.07; I^2=54.8%) demonstrated no statistically significant difference compared with the control groups. Detailed findings are presented in Figures 6-9.

Figure 6. Subgroup analysis by physical activity features [22-29]. ES: effect size; OB: objective; PA: physical activity; SR: self-reported.



XSL•FO
RenderX

Figure 7. Subgroup analysis by intervention category [22-29]. ES: effect size; OB: objective; PA: physical activity; SR: self-reported; WA: walking activity.

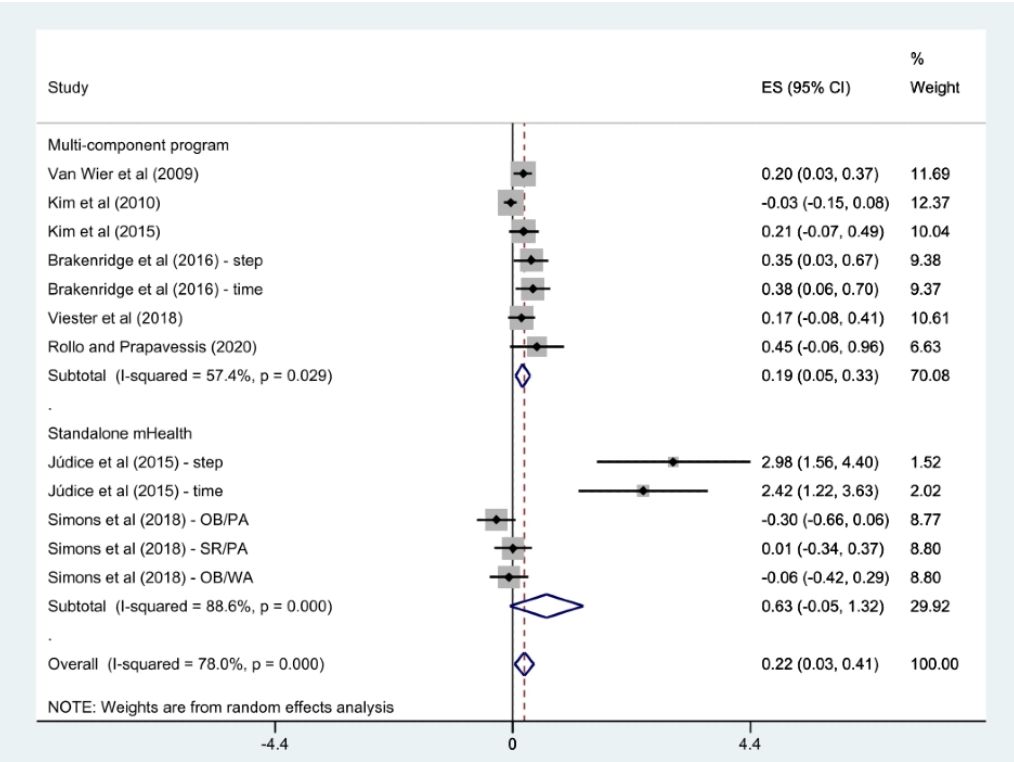
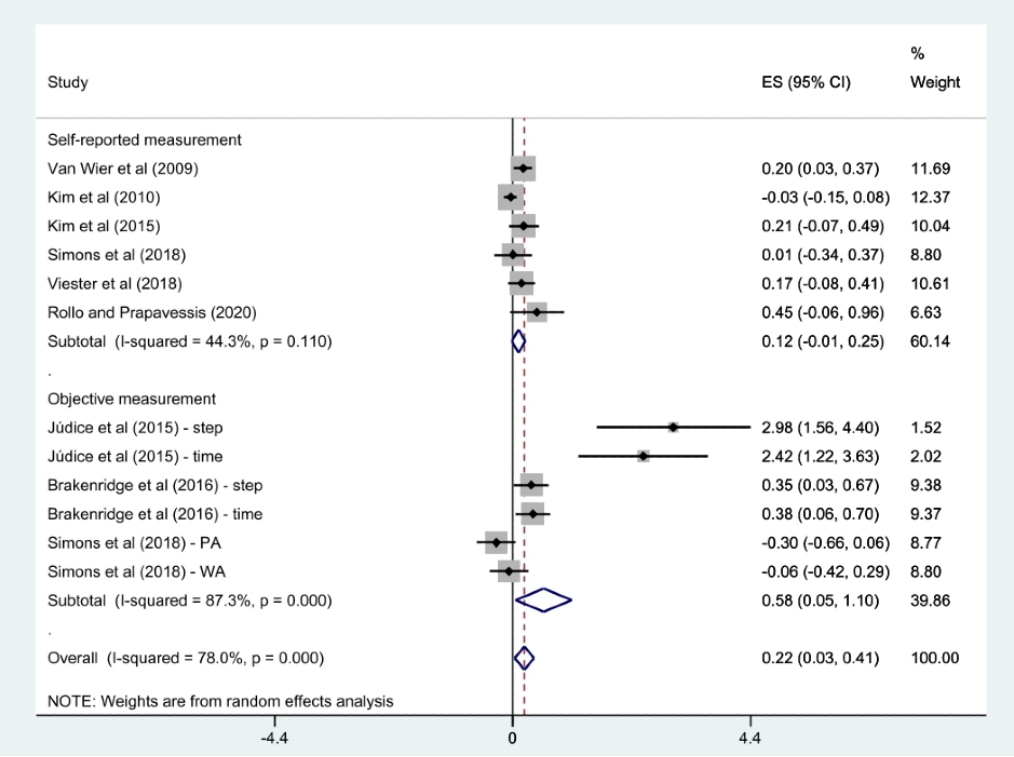
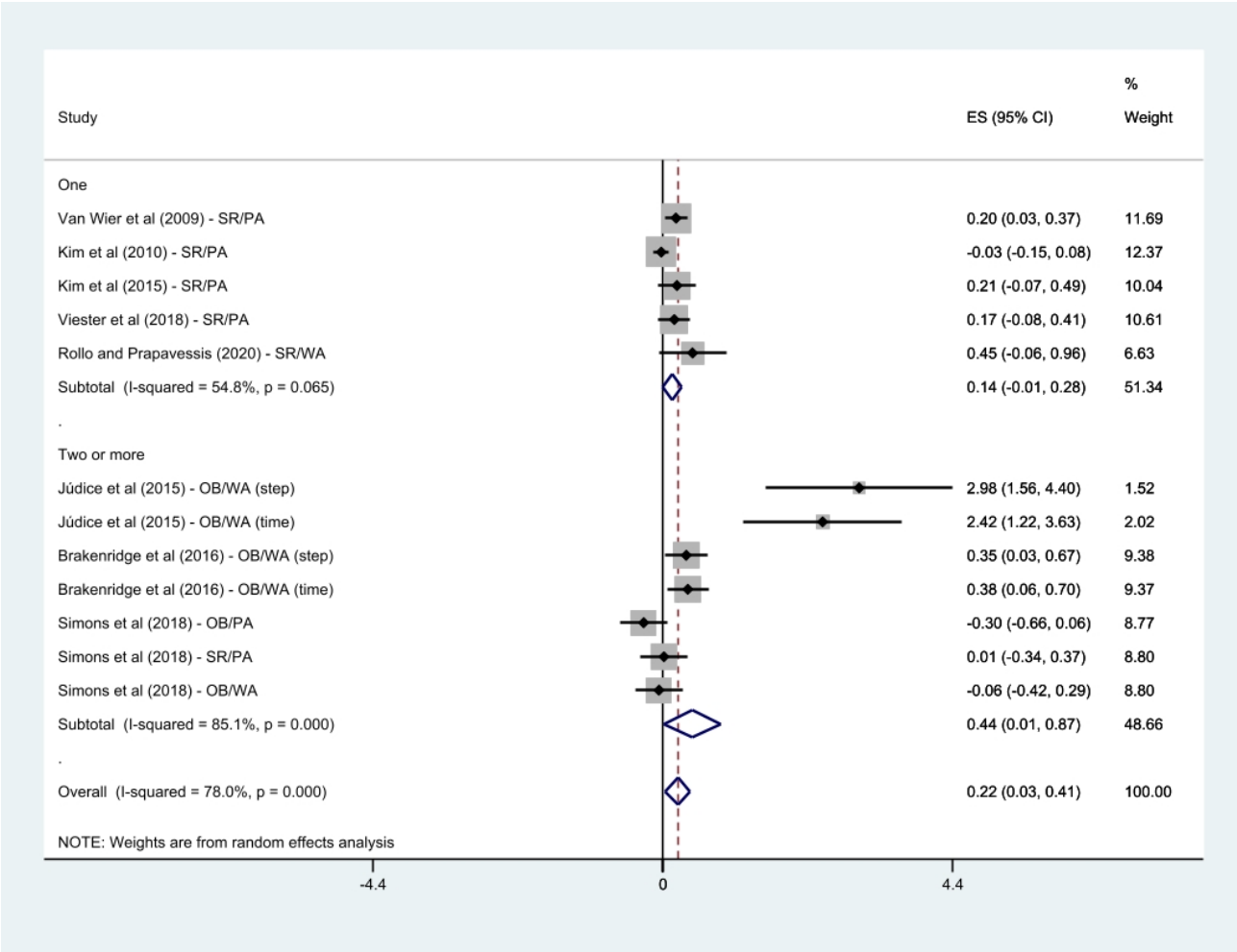


Figure 8. Subgroup analysis by physical activity measurements [22-29]. ES: effect size; PA: physical activity; WA: walking activity.



XSL•FO
RenderX

Figure 9. Subgroup analysis by the number of delivery modes [22-29]. ES: effect size; OB: objective; PA: physical activity; SR: self-reported; WA: walking activity.



Publication Bias and Sensitivity Analysis

Publication bias was assessed using a funnel plot (Figures 10 and 11). Although the funnel plot was shown to be visually asymmetrical for PA, the Begg correlation test (*P*=.15) was not statistically significant.

We conducted a sensitivity analysis to estimate the robustness of our findings (Multimedia Appendices 3 and 4). We identified the weights of the included studies and then eliminated them one by one to assess the impact of the study on the overall effects. When the study by Kim et al [25] was excluded, there was a change in the MD because its weight was the largest in PA analysis (SMD 0.19, 95% CI 0.10-0.29). After removal of other studies, the MD ranged from 0.08 to 0.19 and was similar to SMD 0.10 (95% CI 0.03-0.18) from the original value calculated using a fixed model.

XSL•FO
RenderX

Figure 10. Funnel plot for publication bias assessment (physical activity).

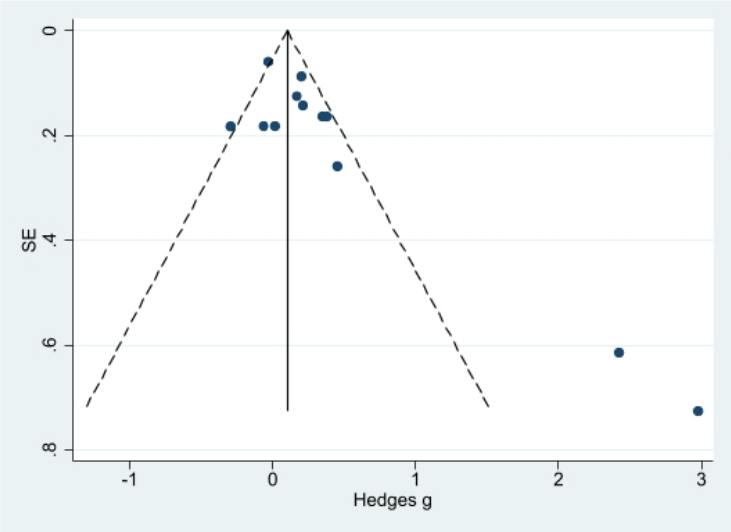
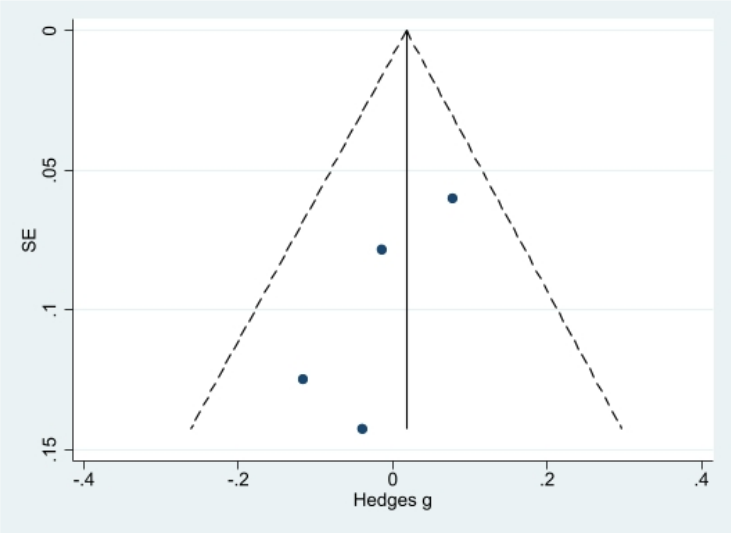


Figure 11. Funnel plot for publication bias assessment (weight loss).



Discussion

Principal Findings

This meta-analysis, which included only RCTs, attempted to analyze the effectiveness of mHealth interventions in PA improvement and weight loss among workers. Subgroup analysis was based on differences in PA features, intervention categories, PA measurement, and the number of delivery modes. Overall, a small to moderate effect was found in mHealth interventions for workers in PA improvement and no statistically significant difference was found in weight loss.

PA improvement was particularly observed in the walking activity feature but not for weight loss in the subgroup analysis. The results also indicated positive effects for multicomponent programs rather than stand-alone mHealth programs in improving overall PA among workers. Moreover, the objective measurement of PA and 2 or more delivery modes were significantly associated with an enhancement in PA when compared to the counterparts.

Limitations

This meta-analysis showed that mHealth interventions could promote PA among the included working populations. However, this study has several limitations that need to be addressed. First, the findings of this study should be interpreted cautiously, considering the relatively small sample sizes and short-term intervention periods (mean 20, SD 14.77 weeks) without follow-up (only 3/8, 38% of the studies reported follow-up). The maintenance of health behavior change is crucial for health promotion practice [40]. Hence, studies with a larger sample size and an extended follow-up period are needed to increase the generalizability of our findings. In addition, including studies with a small sample size may result in errors owing to small-study effects because the effect size might be relatively large. Second, although the heterogeneity was lowered in the subgroup analysis, there was substantial heterogeneity in the main analysis of this study. The possible explanation is that the heterogeneity is because of the additional intervention contents, difference in frequency, intervention duration, and delivery methods. Third, it is considered necessary to compare the differences between PA promotion programs with behavior

XSL•FO
RenderX

change and weight management programs. Unfortunately, in the studies included in our analysis, it was difficult to separate them into 2 distinct classifications. Fourth, our search was restricted only to full-text articles published in English or Korean; thus, language and publication bias might have resulted when relevant studies outside the current scope were excluded. Finally, the outcome variables in this meta-analysis were excluded by using only subjective, self-reported data from the previous studies (5/8, 63%). Although the subjective measurement of PA was performed using validated tools (International PA Questionnaire, Short Questionnaire to Assess Health-Enhancing PA, etc), the potential bias for self-reported data cannot be ignored.

Comparison With Previous Work

Mobile technologies (eg, mobile phones, tablets, and tracking devices) have offered an innovative delivery method for promoting PA in public health practice [41]. Although many scholars have used mHealth interventions as a useful method for behavior change, their effectiveness remains uncertain [35]. Moreover, there is a review deficit for the target populations and settings using mHealth for PA promotion and weight loss. Indeed, there was a literature review that concluded that mHealth interventions for workers are a practical and effective way to promote PA [33]. However, a meta-analysis related to this review [33] was not performed owing to the heterogeneity of the studies' outcomes and methods and incomplete reporting. To our knowledge, this meta-analysis is the first study to examine the effectiveness of mHealth interventions for PA promotion and weight loss in the working population.

Despite a lack of review studies on the working population, various reviews with the general adult population have shown a positive effect of mHealth in promoting PA. These previous studies concluded that interventions comprising wearable devices and smartphone apps effectively promoted PA in adults, with small to moderate effects (SMD 0.43, 95% CI 0.03-0.82; SMD 0.27, 95% CI 0.15-0.39) [42,43]. Similar conclusions were reported in a meta-regression study [44]. Furthermore, Schoeppe et al [45] found significant PA improvement via smartphone apps. However, there were nonsignificant differences in PA observed by Flores et al [46] (SMD 0.40, 95% CI –0.07 to 0.87), Direito et al [20] (SMD 0.14, 95% CI –0.12 to 0.41), and Islam et al [47] (MD 0.17, 95% CI –2.21 to 2.55).

This study showed evident, positive effects for walking activity using subgroup analysis. The finding agrees with the results of study by Tang et al [48], which reported that the use of a wearable tracker was associated with improvements in PA, especially in the number of steps (SMD 0.332, 95% CI 0.16-0.50). Gal et al [42] and Feter et al [49] also reported that interventions using mobile phones have resulted in significant enhancement on the number of steps (SMD 0.51, 95% CI 0.12-0.91; MD 735, 95% CI 28-1243, respectively). However, Romeo et al [50] and Direito et al [20] could not find significant improvements in walking activity (MD 477, 95% CI –230 to 1183; SMD 0.14, 95% CI –0.01 to 0.29, respectively). In this study, a small effect on overall PA and walking activity was observed. Given the heterogeneity of the included studies, the potential effects of promoting overall PA and walking activity by mHealth interventions cannot be ignored.

Following the recommendation of a previous review, mHealth intervention programs for improving PA should focus on participants' weight, waist circumference, and BMI [51]. Hence, we considered weight as a secondary outcome with several mHealth intervention studies among the included populations. Islam et al [47] evaluated the effectiveness of mHealth interventions for weight management and found a small but significant loss. In addition, previous studies reported pooled effects of interventions via smartphone app on weight loss (–1.04 kg, 95% CI –1.75 to –0.34; –2.56 kg, 95% CI –3.46 to –1.65) [46,52]. However, the effect of secondary analysis on weight loss was not statistically significant in this study. A previous review revealed that weight management programs combining PA and diet were more effective than interventions with PA alone for weight loss [53]. We hypothesized that one of the reasons for the inconsistent results for weight loss could be the differences in the interventional focus of the study. Indeed, half of the studies included secondary analysis on weight loss (4/8, 50%) and intervention contents for improving PA and diet. Among these 4 studies, 2 (50%) studies reported significant weight loss [28,29]. We can confirm that the studies that obtained significant weight loss results included an intensive focus on diet behavior compared with other studies. A previous study had also emphasized the importance of dietary change in the weight loss program [53]. Moreover, the variability in participants' characteristics and interventions' intensity, duration, and type in each study could make the results inconsistent. The inclusion of a study with overweight workers is also a probable cause for the different outcomes. Moreover, there was no clear evidence of benefit from interventions with a wearable tracker for weight loss or PA in overweight populations [48].

In addition, it was impossible to draw any definitive conclusions on the relative effectiveness of different delivery methods owing to considerable heterogeneity and the small number of high-quality studies. However, we found evidence that stand-alone mHealth interventions with no additional *offline* components were less likely to increase PA. The evidence supported the results of various studies and could lead to more robust results. Previous reviews have suggested that behavioral and health outcomes of multicomponent interventions are better than those of stand-alone mHealth interventions [45]. Islam et al [47] also reported that various delivery channels are deemed effective for reducing weight and maintaining BMI. The mHealth technologies incorporated into existing programs by educational support were reported to be beneficial [51].

On the other hand, a previous review has pointed out that mHealth devices were mainly used for outcome measurement or as a supplement to other intervention components [54]. In addition, another review suggested that most mHealth interventions support increasing PA levels, especially by using SMS text messaging and facilitating self-monitoring [55]. In this study, most of the included studies used 2 or more mHealth technologies, and some of them used mHealth devices for outcome measurements. On the basis of the mixed results, a

XSL•FO
RenderX

clearer conclusion could be inferred through analysis by including more studies in the future.

Recommendations for Future Research

This study examined the effectiveness of mHealth intervention for PA improvement among workers and found a modest benefit of the studies. The role of behavior change techniques within the mHealth interventions can also affect PA [44,47]. It was impossible to analyze the effectiveness of each behavior change technique as they were mixed with other intervention characteristics. Hence, further research should include research questions on related aspects for the behavior change techniques. In addition, it was difficult to compare the differences between behavior change programs and weight management programs in this analysis. Finally, further studies to assess feasibility, long-term impact with follow-up, and engagement of mHealth interventions are recommended. This study could not estimate which delivery modes were most likely to change behaviors owing to the few high-quality studies and heterogeneity. With more extensive studies, we propose to analyze the effectiveness of the type of delivery mode on behavior changes in future research.

Conclusions

Although the overall effects might be relatively small, mHealth interventions appeared to be effective for improving PA among workers. Multicomponent interventions using mHealth devices were more effective than stand-alone uses of mHealth devices. Future studies, including a larger sample size with extended periods, are required to evaluate the effects of behavior change techniques within mHealth interventions on workers' PA improvement and weight management.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea and funded by the Ministry of Education (grant number: NRF-2019R1I1A1A01060534).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Electronic search strategy in electronic databases.
[PDF File (Adobe PDF File), 208 KB - mhealth_v10i1e30682_app1.pdf]

Multimedia Appendix 2
Program objectives and contents among included studies.
[PDF File (Adobe PDF File), 199 KB - mhealth_v10i1e30682_app2.pdf]

Multimedia Appendix 3
Sensitivity analysis for physical activity.
[PDF File (Adobe PDF File), 133 KB - mhealth_v10i1e30682_app3.pdf]

Multimedia Appendix 4
Sensitivity analysis for weight loss.
[PDF File (Adobe PDF File), 123 KB - mhealth_v10i1e30682_app4.pdf]

References

1. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: a brief review. Sports Med Health Sci 2019 Dec;1(1):3-10. [doi: 10.1016/j.smhs.2019.08.006]
2. Nocon M, Hiemann T, Müller-Riemenschneider F, Thalau F, Roll S, Willich SN. Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. Eur J Cardiovasc Prev Rehabil 2008 Jun;15(3):239-246. [doi: 10.1097/HJR.0b013e3282f55e09] [Medline: 18525377]
3. Diaz KM, Shimbo D. Physical activity and the prevention of hypertension. Curr Hypertens Rep 2013 Dec 20;15(6):659-668 [FREE Full text] [doi: 10.1007/s11906-013-0386-8] [Medline: 24052212]
4. Cleven L, Krell-Roesch J, Nigg CR, Woll A. The association between physical activity with incident obesity, coronary heart disease, diabetes and hypertension in adults: a systematic review of longitudinal studies published after 2012. BMC Public Health 2020 May 19;20(1):726 [FREE Full text] [doi: 10.1186/s12889-020-08715-4] [Medline: 32429951]
5. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. Prev Med 2007 Dec;45(6):401-415. [doi: 10.1016/j.ypmed.2007.07.017] [Medline: 17707498]
6. Global Action Plan on Physical Activity 2018-2030: More Active People for a Healthier World. Geneva: World Health Organization; 2019.

XSL•FO
RenderX

7. Ryu H, Jung J, Moon J. Effectiveness of a mobile health management program with a challenge strategy for improving the cardiovascular health of workers. J Occup Environ Med 2021 Mar 01;63(3):e132-e137. [doi: 10.1097/JOM.0000000000002130] [Medline: 33395169]
8. Labor force participation rate, total (% of total population ages 15+). International Labour Organization. URL: https://data.worldbank.org/indicator/SL.TLF.CACT.ZS [accessed 2022-01-05]
9. Thivel D, Tremblay A, Genin PM, Panahi S, Rivière D, Duclos M. Physical activity, inactivity, and sedentary behaviors: definitions and implications in occupational health. Front Public Health 2018 Oct 5;6:288 [FREE Full text] [doi: 10.3389/fpubh.2018.00288] [Medline: 30345266]
10. de Korte E, Wiezer N, Bakhuys Roozeboom M, Vink P, Kraaij W. Behavior change techniques in mHealth apps for the mental and physical health of employees: systematic assessment. JMIR Mhealth Uhealth 2018 Oct 03;6(10):e167 [FREE Full text] [doi: 10.2196/mhealth.6363] [Medline: 30282621]
11. Azevedo LM, Chiavegato LD, Carvalho CR, Braz JR, Nunes Cabral CM, Padula RS. Are blue-collar workers more physically active than white-collar at work? Arch Environ Occup Health 2021;76(6):338-347. [doi: 10.1080/19338244.2020.1835796] [Medline: 33092496]
12. Michaels CN, Greene AM. Worksite wellness: increasing adoption of workplace health promotion programs. Health Promot Pract 2013 Jul;14(4):473-479. [doi: 10.1177/1524839913480800] [Medline: 23545334]
13. Arena R, Guazzi M, Briggs PD, Cahalin LP, Myers J, Kaminsky LA, et al. Promoting health and wellness in the workplace: a unique opportunity to establish primary and extended secondary cardiovascular risk reduction programs. Mayo Clin Proc 2013 Jun;88(6):605-617 [FREE Full text] [doi: 10.1016/j.mayocp.2013.03.002] [Medline: 23726400]
14. Hendriksen I, Snoijer M, de Kok BP, van Vilsteren J, Hofstetter H. Effectiveness of a multilevel workplace health promotion program on vitality, health, and work-related outcomes. J Occup Environ Med 2016 Jun;58(6):575-583 [FREE Full text] [doi: 10.1097/JOM.0000000000000747] [Medline: 27136605]
15. Abraham C, Graham-Rowe E. Are worksite interventions effective in increasing physical activity? A systematic review and meta-analysis. Health Psychol Rev 2009 Mar;3(1):108-144. [doi: 10.1080/17437190903151096]
16. de Korte EM, Wiezer N, Janssen JH, Vink P, Kraaij W. Evaluating an mHealth app for health and well-being at work: mixed-method qualitative study. JMIR Mhealth Uhealth 2018 Mar 28;6(3):e72 [FREE Full text] [doi: 10.2196/mhealth.6335] [Medline: 29592846]
17. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
18. Nacinovich M. Defining mHealth. J Commun Healthcare 2013 Jul 18;4(1):1-3 [FREE Full text] [doi: 10.1179/175380611X12950033990296]
19. mHealth. In: Information and Communications for Development. Washington D.C: World Bank Group; 2012.
20. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
21. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
22. Brakenridge CL, Fjeldsoe BS, Young DC, Winkler EA, Dunstan DW, Straker LM, et al. Evaluating the effectiveness of organisational-level strategies with or without an activity tracker to reduce office workers' sitting time: a cluster-randomised trial. Int J Behav Nutr Phys Act 2016 Nov 04;13(1):115 [FREE Full text] [doi: 10.1186/s12966-016-0441-3] [Medline: 27814738]
23. Júdice PB, Hamilton MT, Sardinha LB, Silva AM. Randomized controlled pilot of an intervention to reduce and break-up overweight/obese adults' overall sitting-time. Trials 2015 Nov 02;16:490 [FREE Full text] [doi: 10.1186/s13063-015-1015-4] [Medline: 26525049]
24. Kim J, Oh S, Steinhubl S, Kim S, Bae WK, Han JS, et al. Effectiveness of 6 months of tailored text message reminders for obese male participants in a worksite weight loss program: randomized controlled trial. JMIR Mhealth Uhealth 2015 Feb 03;3(1):e14 [FREE Full text] [doi: 10.2196/mhealth.3949] [Medline: 25648325]
25. Kim Y, Pike J, Adams H, Cross D, Doyle C, Foreyt J. Telephone intervention promoting weight-related health behaviors. Prev Med 2010 Mar;50(3):112-117. [doi: 10.1016/j.ypmed.2009.12.008] [Medline: 20006642]
26. Rollo S, Prapavessis H. A combined health action process approach and mHealth intervention to increase non-sedentary behaviours in office-working adults-a randomised controlled trial. Appl Psychol Health Well Being 2020 Nov;12(3):660-686. [doi: 10.1111/aphw.12201] [Medline: 32342662]
27. Simons D, De Bourdeaudhuij I, Clarys P, De Cocker K, Vandelanotte C, Deforche B. Effect and process evaluation of a smartphone app to promote an active lifestyle in lower educated working young adults: cluster randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 24;6(8):e10003 [FREE Full text] [doi: 10.2196/10003] [Medline: 30143477]

XSL•FO
RenderX

28. van Wier MF, Ariëns GA, Dekkers JC, Hendriksen IJ, Smid T, van Mechelen W. Phone and e-mail counselling are effective for weight management in an overweight working population: a randomized controlled trial. BMC Public Health 2009 Jan 09;9:6 [FREE Full text] [doi: 10.1186/1471-2458-9-6] [Medline: 19134171]
29. Viester L, Verhagen EA, Bongers PM, van der Beek AJ. Effectiveness of a worksite intervention for male construction workers on dietary and physical activity behaviors, body mass index, and health outcomes: results of a randomized controlled trial. Am J Health Promot 2018 Mar;32(3):795-805 [FREE Full text] [doi: 10.1177/0890117117694450] [Medline: 28730841]
30. Haque MS, Kangas M, Jämsä T. A persuasive mHealth behavioral change intervention for promoting physical activity in the workplace: feasibility randomized controlled trial. JMIR Form Res 2020 May 04;4(5):e15083 [FREE Full text] [doi: 10.2196/15083] [Medline: 32364506]
31. Ganesan AN, Louise J, Horsfall M, Bilsborough SA, Hendriks J, McGavigan AD, et al. International mobile-health intervention on physical activity, sitting, and weight: the Stepathlon cardiovascular health study. J Am Coll Cardiol 2016 May 31;67(21):2453-2463 [FREE Full text] [doi: 10.1016/j.jacc.2016.03.472] [Medline: 27050185]
32. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: a qualitative exploration of the role of healthcare partnerships and mHealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016 Jan 20;4(1):e5 [FREE Full text] [doi: 10.2196/mhealth.4814] [Medline: 26792225]
33. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit Health 2019;5:2055207619839883 [FREE Full text] [doi: 10.1177/2055207619839883] [Medline: 30944728]
34. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
35. Gilson ND, Pavey TG, Wright OR, Vandelanotte C, Duncan MJ, Gomersall S, et al. The impact of an m-Health financial incentives program on the physical activity and diet of Australian truck drivers. BMC Public Health 2017 May 18;17(1):467 [FREE Full text] [doi: 10.1186/s12889-017-4380-y] [Medline: 28521767]
36. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
37. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
38. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
39. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
40. Orleans CT. Promoting the maintenance of health behavior change: recommendations for the next generation of research and practice. Health Psychol 2000 Jan;19(1S):76-83. [doi: 10.1037/0278-6133.19.suppl1.76] [Medline: 10709951]
41. mHealth: New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
42. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 03;4(1):42 [FREE Full text] [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
43. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
44. Eckerstorfer LV, Tanzer NK, Vogrincic-Haselbacher C, Kedia G, Brohmer H, Dinslaken I, et al. Key elements of mHealth interventions to successfully increase physical activity: meta-regression. JMIR Mhealth Uhealth 2018 Nov 12;6(11):e10076 [FREE Full text] [doi: 10.2196/10076] [Medline: 30425028]
45. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
46. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
47. Islam MM, Poly TN, Walther BA, Jack Li YC. Use of mobile phone app interventions to promote weight loss: meta-analysis. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e17039 [FREE Full text] [doi: 10.2196/17039] [Medline: 32706724]
48. Tang MS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e15576 [FREE Full text] [doi: 10.2196/15576] [Medline: 32706685]

XSL•FO
RenderX

49. Feter N, Dos Santos TS, Caputo EL, da Silva MC. What is the role of smartphones on physical activity promotion? A systematic review and meta-analysis. Int J Public Health 2019 Jun;64(5):679-690. [doi: 10.1007/s00038-019-01210-7] [Medline: 30758514]
50. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
51. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
52. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28 Suppl 2:1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
53. Johns DJ, Hartmann-Boyce J, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Diet or exercise interventions vs combined behavioral weight management programs: a systematic review and meta-analysis of direct comparisons. J Acad Nutr Diet 2014 Oct;114(10):1557-1568 [FREE Full text] [doi: 10.1016/j.jand.2014.07.005] [Medline: 25257365]
54. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
55. O'Reilly GA, Spruijt-Metz D. Current mHealth technologies for physical activity assessment and promotion. Am J Prev Med 2013 Oct;45(4):501-507 [FREE Full text] [doi: 10.1016/j.amepre.2013.05.012] [Medline: 24050427]

Abbreviations

MD: mean difference
mHealth: mobile health
PA: physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMD: standardized mean difference

Edited by L Buis; submitted 25.05.21; peer-reviewed by S Buckingham, K Basen-Engquist; comments to author 26.07.21; revised version received 17.09.21; accepted 14.12.21; published 21.01.22.

Please cite as:
Jung J, Cho I
Promoting Physical Activity and Weight Loss With mHealth Interventions Among Workers: Systematic Review and Meta-analysis of Randomized Controlled Trials
JMIR Mhealth Uhealth 2022;10(1):e30682
URL: https://mhealth.jmir.org/2022/1/e30682
doi:10.2196/30682
PMID:35060913

©Jiyeon Jung, Inhae Cho. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 21.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Use of Mobile Apps for Self-care in People With Parkinson Disease: Systematic Review

JuHee Lee[1,2], PhD, GNP-BC, RN; Insun Yeom[2], PhD, RN; Misook L Chung[3,4], PhD, RN; Yielin Kim[4], MSN, CRRN, RN; Subin Yoo[2], RN; Eunyoung Kim[2], MSN, RN

[1]Mo-Im Kim Nursing Research Institute, Yonsei Evidence Based Nursing Centre of Korea: A JBI Affiliated Group, College of Nursing, Yonsei University, Seoul, Republic of Korea

[2]Brain Korea 21 FOUR Project, College of Nursing, Yonsei University, Seoul, Republic of Korea

[3]College of Nursing, University of Kentucky, Lexington, KY, United States

[4]College of Nursing, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Eunyoung Kim, MSN, RN
Brain Korea 21 FOUR Project
College of Nursing
Yonsei University
College of Nursing, Yonsei University
50-1 Yonsei-ro, Seodaemoon-gu
Seoul, 03722
Republic of Korea
Phone: 82 10 2816 2884
Fax: 82 2 2227 8303
Email: key7481@daum.net

Abstract

Background: Self-care is essential for people with Parkinson disease (PD) to minimize their disability and adapt to alterations in physical abilities due to this progressive neurodegenerative disorder. With rapid developments in mobile technology, many health-related mobile apps for PD have been developed and used. However, research on mobile app–based self-care in PD is insufficient.

Objective: This study aimed to explore the features and characteristics of mobile apps for self-care in people with PD.

Methods: This study was performed sequentially according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement. PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature, Cochrane Library, Web of Science, and PsycINFO were searched in consultation with a librarian on June 8, 2021. We used keywords including ”Parkinson disease” and ”mobile.”

Results: A total of 17 studies were selected based on the inclusion criteria, including 3 randomized controlled trials and 14 observational studies or quasi-experimental studies. The use of mobile apps for self-care in people with PD focused on symptom monitoring, especially motor symptoms. Motor symptoms were objectively measured mainly through the sensors of smartphones or wearable devices and task performance. Nonmotor symptoms were monitored through task performance or self-reported questionnaires in mobile apps. Most existing studies have focused on clinical symptom assessment in people with PD, and there is a lack of studies focusing on symptom management.

Conclusions: Mobile apps for people with PD have been developed and used, but strategies for self-management are insufficient. We recommend the development of mobile apps focused on self-care that can enhance symptom management and health promotion practices. Studies should also evaluate the effects of mobile apps on symptom improvement and quality of life in people with PD.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42021267374; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021267374.

(JMIR Mhealth Uhealth 2022;10(1):e33944) doi:10.2196/33944

XSL•FO
RenderX

KEYWORDS

systematic review; Parkinson disease; motor symptoms; nonmotor symptoms; smartphone; mobile phone; mobile health; mobile apps; self-care; symptom; monitoring; review; disability; app; care; quality of life; self-management

Introduction

The number of people with Parkinson disease (PD) has increased significantly with the aging population and rising life expectancy [1]. According to a systematic literature review that analyzed 47 studies, PD is predominantly prevalent in older adults (aged above 70 years) [2]. A study estimating life years and the prevalence of PD from 1990 to 2016 reported that the worldwide burden related to PD had more than doubled [1].

People with PD experience motor and nonmotor symptoms. Most motor symptoms include tremors, postural instability, bradykinesia, and rigidity. Nonmotor symptoms are associated with sensory abnormalities, neuropsychiatric abnormalities, sleep disorders, and autonomic dysfunction (eg, bladder, bowel, and sexual dysfunction) [3,4]. Symptom management is essential to maintain one's functional ability, as insufficiently managed PD symptoms negatively influence quality of life and worsen physical disabilities in people with PD [5]. As defined by the theory of self-care in chronic illness, self-care in individuals with chronic diseases refers to a series of processes for maintaining health [6]. This self-care process includes detecting, interpreting, and responding to altered symptoms [6]. For effective self-care, symptom monitoring is essential to recognize changes in symptoms, along with skills to manage symptoms and perform health promotion practices [6].

Traditional interventions to improve self-care in PD have used face-to-face instruction to deliver health-promoting information, rehabilitation therapy, or interventions aiming to induce cognitive behavioral changes. Previous review studies on self-care interventions in people with PD identified interventions, most of which involved self-care management or self-care maintenance (eg, exercise, occupational therapy, health coaching, psychological strategy training, and lifestyle advice) to improve patients' health outcomes [7,8]. All these were face-to-face interventions delivered without using mobile technology.

Mobile health (mHealth) devices have enabled improvements in diagnosis and treatment, as well as connection with distant patients [9]. Over the past few decades, dramatic advances in computer and communication technologies have led to the development of mHealth and communication technologies in the medical environment [10]. The portability and wide distribution of smartphones have enabled the development and usage of various health care apps that can track and manage symptoms, and these have strengthened self-care interventions for people with chronic illness. For example, recent systematic reviews have reported that mobile apps for type 2 diabetes that provide goal management or motivational feedback based on self-reported symptoms or vital sign monitoring are effective in reducing the fasting blood sugar and waist circumference [11,12]. In addition, a study reported that the overall survival rate of patients with advanced lung cancer improved after implementing a tracking algorithm, referred to as an "e-follow-up application," via early relapse detection using weekly self-reports of symptoms [13].

Many mobile apps for PD patients have been developed and implemented. Moreover, 2 systematic reviews focusing on apps available in Google Play and the App Store from 2011 to 2016 found 92 and 125 apps, respectively, that were potentially useful for individuals with PD [14,15]. These reviews were conducted to identify a suitable operating system for these apps and analyze their usability and validity. However, both reviews did not provide detailed analyses regarding the use of mobile apps in self-care interventions. As there is no available curative treatment for PD, the severity of the symptoms and disease should be closely monitored to manage PD effectively. Symptom tracking using a smartphone offers the possibility of regularly monitoring patients' symptoms over time, thereby overcoming the problem with traditional clinical assessments that provide a "snapshot" of patients' conditions [16].

This study was performed to explore the use of mobile apps for self-care in people with PD. We specifically explored the features and characteristics of the mobile apps that were used for self-care maintenance, self-care monitoring, and self-care management.

Methods

Design

This study is a systematic review following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [17]. The protocol was registered in the International Prospective Register of Systematic Reviews (Trial registration number: CRD42021267374).

Search Strategy

The literature search was conducted in 3 steps. First, a search was conducted in PubMed using the following relevant MeSH (Medical Subject Headings) terms and free-text keywords. The term "Parkinson disease" and "mobile" were used as the keywords for the concept, and MeSH or Emtree terms linked to the search domains were used. The final search query was developed in consultation with a librarian having a PhD degree and more than 10 years of experience (see Multimedia Appendix 1). In the second step, a literature search was conducted in PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature, Cochrane Library, Web of Science, and PsycINFO using the search query on June 8, 2021. All search results were reviewed by the librarian. In the last step, the references of the selected studies were manually searched by 2 researchers.

Eligibility Criteria for the Review

The studies for the review were restricted to those related to self-care using mobile apps in adults with PD. We also included studies that were published in English from January 2003 to June 2021 in peer-reviewed journals. This start date was chosen

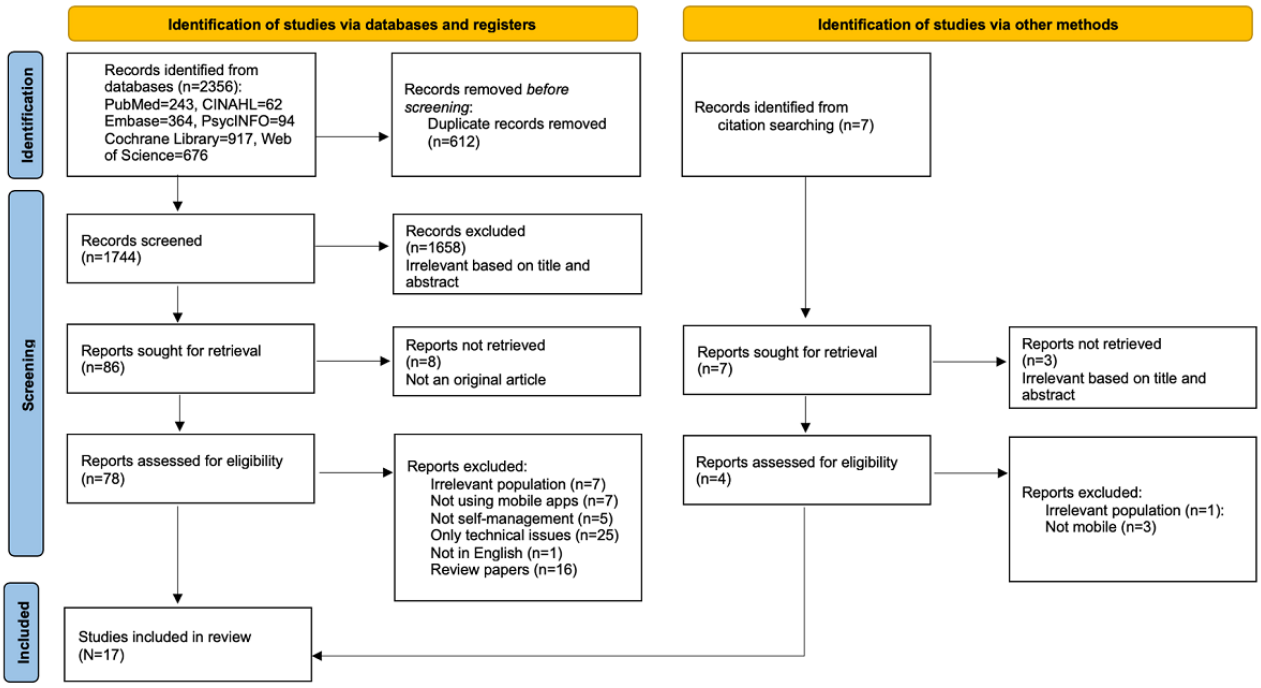
XSL•FO
RenderX

because terms referring to phenomena such as cell phones, computers, handheld devices, and small portable wireless devices were introduced in 2003 as MeSH terms. In this study, self-care is defined as health maintenance practices, symptom tracking and monitoring, and management of symptoms [6]. Mobile apps are generally defined as computer programs or software applications for a mobile device such as a smartphone. We excluded studies that evaluated only technical issues related to mobile apps or tested them with healthy adults or those with other chronic diseases.

Study Selection

All the study selection steps were initially performed by 2 researchers (EK and YK). We identified a total of 2356 studies from all databases searched in the initial stage and removed 612 duplicates. The titles and abstracts of all the remaining 1744 records were screened for potential relevance based on a standardized checklist. Of those studies, 1658 were excluded because they were considered irrelevant to the purpose of this study. In addition, 8 studies were excluded because they were not original articles, and following a full-text review, 61 studies were excluded. The reasons for exclusion were that the population did not meet the inclusion criteria, a mobile app was not used, there was no self-care context, the articles dealt with only technical issues, or they were review articles. Citation searching yielded 7 documents that were excluded as irrelevant through title, abstract, and full-text assessment. Finally, 17 studies were selected for this review, as shown in Figure 1.

Figure 1. Flow diagram of the search.



Data Extraction, Analysis, and Synthesis

Data extraction was performed independently by 2 researchers (EK and SY) using a standardized format. The following data were extracted: author(s); published year; title; published journal; country where the study was performed; aim of the study; design of the study; participants' characteristics; the name of the mobile app; and the intervention duration, results, and limitations. For data analysis, the type of mobile app was categorized based on the method of symptom data collection and other functions. The outcome measure was categorized as satisfaction with the app, feasibility, symptom severity, and patient outcomes. The characteristics of the mobile apps were classified as self-care maintenance, self-care monitoring, and self-care management based on the theory of self-care in chronic illness [6]. Self-care maintenance was defined as health-promoting practices to maintain good health status, such as physical activity, treatment adherence, a regular sleep pattern, and nutritional intake whereas self-care monitoring was defined as tracking and recognizing symptoms leading to interpretation. Symptom monitoring was divided into monitoring of motor and nonmotor symptoms, and each symptom was classified with reference to the literature [3,4]. Self-care management pertained to behavioral changes, such as changes in the activity level, medication use, information seeking, and dietary changes. Self-care management requires symptom recognition and interpretation when physical changes occur.

Quality Appraisal

The quality of the selected studies was assessed using tools for assessing risk of bias developed by the Cochrane Collaboration. The risk of bias in non-randomized studies of interventions (ROBINS-I) [18] was used for quality assessment of observational studies and quasi-experimental studies. The revised Cochrane risk of bias tool for randomized trials (RoB2) [19] was used for randomized controlled trials (RCTs). ROBINS-I evaluates the risk of bias in the confounding variables, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, selection of the reported results, and overall bias. Each section is evaluated as low, moderate, serious, critical, and no information. RoB2 consists of 6 sections,

XSL•FO
RenderX

including the randomization process, deviation from the intended interventions, missing outcome data, measurement of the outcome, selection of the results, and the risk of overall bias. In each section, the risk of bias is evaluated using questions with responses “yes,” “probably yes,” “no,” “probably no,” and “no information,” and each section is finally judged as low risk, some concerns, or high risk according to the evaluation algorithm. The quality of the studies was assessed independently by 2 researchers (IY and EK). Any discrepancies were resolved by consensus.

Results

Study Characteristics

In total, 17 articles were analyzed in this study, as shown in Table 1. Publication years ranged from 2013 to 2020. Of the 17 selected studies, 6 were published in 2020 (35.3%). There were 12 observational studies (70.6%), 2 quasi-experimental studies (11.8%), and 3 RCTs (17.6%). The study of Gatsios et al [20] was classified as an observational study because it analyzed only the intervention group as an ancillary study of an RCT. The intervention duration varied from a single session for 30 minutes [16] to over 6 months [21-23]. More than half of the studies had intervention periods of less than 1 month [16,20,24-29]. We found that 4 studies were conducted through international collaborations in multiple countries [20,24,30,31]. Researchers in the United States conducted 7 studies, followed by England, Finland, Italy, Netherlands, and the United Kingdom with 2 studies each. Further, 1 study each was conducted in Australia, Belgium, Greece, Israel, and Scotland. A total of 1246 people with PD participated in the 17 studies. The participants’ age ranged from 34 to 84 years (mean age=63.02 years), and 58.8% (733) of the participants were male.

XSL•FO
RenderX

Table 1. Characteristics of the included studies.

Author (year)/country	Aim and study design	Participant characteristics (sample size, gender, age, disease duration)	App name	Frequency and duration	Results
Keränen and Liikkanen [32] (2013)/Finland	To evaluate the feasibility of medication reminders SMS; Observational	Total: 45 Male: 29 (64.4%) Age: 66.4 (SD 7.90) y	Not mentioned	4 weeks	Most were satisfied with usability (69%). The majority wanted to continue using the system (80%).
Pan et al [28] (2015)/United States	To develop and test a mobile app to assess motor symptom severity; Observational	Total: 40 Male: 35 (87.5%) Age: 68.5 (SD 9.5) y Disease duration: 6.6 (SD 9.5) y	PD[a] Dr	A single motor performance test session	PD Dr could effectively detect hand resting tremor and gait difficulty and estimate motor symptom severity using the captured motion features.
Kassavetis et al [16] (2015)/United Kingdom	To develop and test stand-alone software for smartphones to assess motor symptoms in PD patients; Observational	Total: 14 Male: 7 (50%) Age: 54.7 (range 34-75) y Disease duration: 3.7 (SD 2.0) y	Not mentioned	A single motor performance test session for 30 minutes	Symptom severity could be assessed from the motion data (tremor, bradykinesia).
Lee et al [29] (2016)/Australia	To generate a predictive model for motor symptom severity using captured data and to evaluate compliance and user satisfaction in a smartphone app; Observational	Total: 103 Male: 52 (50.5%) Age: 66.5 (range 38-91) y Disease duration: 8.75 (range 0.5-24) y	Not mentioned	Twice within 2 weeks	Symptom severity could be assessed from the motion data (tremor, bradykinesia, cognition). A prediction model accounted for 52.3% of the variation in motor symptoms. Participants showed high compliance (96%). Most are satisfied with usability (83%) and usefulness (97%).
Silva de Lima et al [33] (2018)/Netherlands	To assess the relationship between the severity of motor fluctuation and walking time collected using a mobile app; Observational	Total: 304 Male: 164 (54%) Age: 63.1 (SD 8.5) y Disease duration: 6.1 (SD 4.3) y	The Fox Wearable Companion app	24 hours for 13 weeks	Mean walking time was related to the severity of motor symptoms. The postmedication activity was on average higher than the premedication activity.
Zhan et al [21] (2018)/United States	To develop an objective measurement tool (mPDS[b]) to assess PD severity; Observational	Total: 169 (129 PD, 23 clinics with PD, 17 clinics without PD) Age: 58.7 (SD 8.6), 64.6 (SD 11.5), and 54.2 (SD 16.5) y Disease duration: 4.3 (SD 4.4) y, 7.0 (SD 4.1) y, and N/A[c]	HopkinsPD	3 times for 6 months	For mPDS generation, 5 activities were selected (gait, balance, finger tapping, voice, and reaction time). The mPDS detected intraday symptom fluctuations. Motor symptom severity could be estimated from mPDS.
Elm et al [22] (2019)/United States	To evaluate the feasibility of a clinician dashboard to monitor patient symptoms through data collected from ePROs[c] and a smart watch; Observational	Total: 39 Male: 29 (74%) Age: 61.9 (SD 10.5) y Disease duration: 7.1 (SD 4.8) y	Fox Wearable Companion app	3 times for 6 months	Participants' compliance rate was 66%. Medication compliance and the severity of ePRO symptoms from the dashboard were the most beneficial components for clinicians' decisions.
Gatsios et al [20] (2020)/Italy, Greece, England	To evaluate the validity and clinical usefulness of data collected using a smartphone and wearable device; Observational	Total: 75 Male: 43 (60%) Age: 67.7 (SD 8.7) y Disease duration: 9.2 (SD 4.4) y	PD manager	12 hours for 11-14 days	Participants' compliance rate was 87%. Collected data from PD manager effectively detected the tremor.
Habets et al [26] (2020)/Netherlands	To evaluate the validity of the eDiary app to collect data using the EMA[d] method; Observational	Total: 20 Male: 16 (80%) Age: 63 (SD 7) y Disease duration: 8 (SD 6) y	Not mentioned	7 times per day for 14 days	eDiary using EMA effectively captured the relationship between affect, motor performance, and motor symptoms.

XSL•FO
RenderX

Author (year)/country	Aim and study design	Participant characteristics (sample size, gender, age, disease duration)	App name	Frequency and duration	Results
Landers and Ellis [34](2020)/United States	To explore the feasibility, safety, and effectiveness of an exercise program to promote physical activity using a mobile app; Observational	Total: 28 Male: 6 (21.4%) Age: 62.1 (SD 9.6) y Disease duration: 3.3 (SD 2.5) y	9zest Parkinson's Therapy	30-60 minutes, 3-5 times per week for 12 weeks for at least 150 minutes per week	Complete compliance was found in 42.9% of participants, and a majority were satisfied with the app exercise (89.5%). Significant improvement was observed in the PDQ8[e] scores, TUG test[f], and STS test[g] after 8 weeks.
Motolese et al [25] (2020)/Italy	To evaluate the feasibility of remote patient monitoring using a smartphone; Observational	Total: 54 Male: 36 (67%) Age: 66.5 (range 59.7-72.2) y Disease duration: 6.5 (range 4-11) y	EncephaLog Home	At least 2 times per week for 3 weeks	Completed compliance was 29.6%. Motor symptom severity could be estimated from the captured motion data (gait, tapping, tremor, and cognition).
Wu and Cronin-Golomb [27] (2020)/United States	To investigate the relationship between sleep quality and daytime functioning based on data collected using EMA and actigraphy; Observational	Total: 20 Male: 13 (65%) Age: 66.5 (SD 9.3) y Disease duration: 6.0 (SD 4.3) y	SymTrend	Every day over 2 weeks	The compliance rate was 91%-94%. Subjective sleep quality significantly predicted next-day anxiety. Other variables were not related to each other.
Horin et al [35](2019)/United States	To evaluate the usability of a mobile app to improve motor symptoms (gait, speech, and dexterity); Quasi-experimental	Total: 37 (I[h]: 17, C[i]: 20) Male: 22 (60%, I), 26 (70%, C) Age: 63.4 (SD 8.6) y (I), 64.9 (SD 8.4) y (C) Disease duration: 6.7 (SD 5.6) y (I), 6.0 (SD 4.3) y (C)	Beats Medical Parkinson's Treatment App	30-60 minutes, once a day for 90 days	Compliance was moderate (64.6%-67.4%). There were no significant improvements in gait, speech, or dexterity.
Kuosmanen et al [24] (2020)/Finland, United Kingdom	To monitor and evaluate hand tremors using a smartphone game and assess medication effects on hand tremors; Quasi-experimental	Total: 13 Male: 5 (38.5%) Age: 64.7 (SD 6.8) y Disease duration: 7.1 (range 2-17) y	STOP (the Sentient Tracking of Parkinson's) app	For 1 month	Motor symptom severity was estimated from the collected tremor data. Through the collected accelerometer signals, the medication effect on rigidity and bradykinesia was confirmed.
Ginis et al [31] (2016)/Belgium, Israel	To compare the effects of gait training using a mobile app and conventional home-based training; RCT[j] (pilot)	Total: 38 (I: 22, C: 18) Male: 6 (15%, I), 11 (27.8%, C) Age: 67.3 (SD 8.1) y (I), 66.1 (SD 8.1) y (C) Disease duration: 10.7 (SD 5.4) y (I), 11.7 (SD 7.6) y (C)	CuPiD system	30 minutes, at least 3 times per week for 6 weeks, with weekly home visits by the researcher	Both groups showed significant improvements in gait speed. The CuPiD group improved significantly more in balance than the control group.
Lakshminarayana et al [30] (2017)/England, Scotland	To evaluate the effectiveness of mobile apps in monitoring PD symptoms; RCT	Total: 201 (I: 94, C: 107) Male: 128 (63.8%, I), 116 (57.9%, C) Age: 59.9 (SD 9.2) y (I), 60.7 (SD 10.3) y (C) Disease duration: 5.7 (SD 4.2) y (I), 5.5 (SD 4.9) y (C)	PTA (the Parkinson's Tracker App)	Once per day or every other day for 16 weeks	The PTA group reported an improvement in medication adherence and PCQ-PD[k] compared with TAU[l].

XSL•FO
RenderX

Author (year)/country	Aim and study design	Participant characteristics (sample size, gender, age, disease duration)	App name	Frequency and duration	Results
Ellis et al [23] (2019)/United States	To evaluate the safety and effectiveness of an exercise program using the mobile app; RCT (single-blind, pilot)	Total: 44 (I: 23, C: 21) Male: 25 (57.7%, I), 23 (52%, C) Age: 64.8 (SD 8.5) y (I), 63.3 (SD 10.6) y (C) Disease duration: 5.9 (SD 3.5) y (I), 3.7 (SD 2.1) y (C)	Wellpepper	5-7 times or at least 3 times per week for 6 months and later extended to 12 months	Daily steps and 6MWT[m] did not show statistically significant between-group differences. PDQ-39[n] improved in the mobile app group.

[a]PD: Parkinson disease.
[b]mPDS: mobile Parkinson disease score.
[c]N/A: not available.
[d]EMA: ecological momentary assessment.
[e]PDQ8: Parkinson Disease Questionnaire 8.
[f]TUG test: timed up-and-go test.
[g]STS test: sit-to-stand test.
[h]I: intervention group.
[i]C: control group.
[j]RCT: randomized controlled trial.
[k]PCQ-PD: Patient-Centered Questionnaire for Parkinson Disease.
[l]TAU: treatment as usual.
[m]6MWT: 6-meter walking test.
[n]PDQ-39: Parkinson Disease Quality of Life.

Quality Appraisal

The quality appraisal results of the 17 selected studies are as follows. In 14 observational studies and quasi-experimental studies, there was no high risk of bias in terms of the confounding variables, classification of interventions, deviations from intended interventions, missing data, or measurement of outcomes. Among the 14 studies, 1 was evaluated as having “serious” concerns regarding the selection of participants and “critical” concerns for the selection of the reported results [21]. Furthermore, 2 studies were evaluated as having “serious” concerns regarding the selection of participants and the reported results [24,35]. Thus, these 3 studies were evaluated as having “serious” or “critical” concerns in at least 1 of the 7 domains in ROBINS-I, as observed in Table 2. This review was conducted to explore the use of mobile apps in PD and focus on the features and characteristics of these apps, and not to evaluate the effectiveness of interventions. Therefore, 3 studies evaluated as “critical” and “serious” were included in the analysis to determine the usage characteristics of the mobile apps.

RoB2 was used to appraise 3 RCTs of which 2 reported only the baseline characteristics of participants without a prior homogeneity analysis between the intervention and control groups [23,31]. However, these studies reported a computer-generated stratified randomization procedure in the randomization process. Therefore, they were considered as having “low risk” in the randomization process and “low risk” in all the other domains of RoB2. The other study was also deemed to be “low risk” in all the domains of RoB2 [30]. All RCTs were evaluated as having a low risk of bias, as observed in Table 3.

XSL•FO
RenderX

Table 2. Quality appraisal of the studies: risk of bias in nonrandomized studies of interventions.

Study (year)	Confounding	Participant selection	Intervention classification	Deviations from intended interventions	Missing data	Outcome measurements	Selection of the reported results	Overall
Keränen and Liikkanen [32] (2013)	Low	Moderate	Low	Low	Low	Low	Low	Moderate
Pan et al [28] (2015)	Low	Moderate	Moderate	Low	Low	Low	Low	Moderate
Kassavetis et al [16] (2015)	Low	Low	Low	Low	Low	Low	Low	Low
Lee et al [29] (2016)	Low	Low	Moderate	Low	Low	Low	Low	Moderate
Silva de Lima et al [33] (2018)	Low	Low	Low	Low	Low	Low	Low	Low
Zhan et al [21] (2018)	Low	Serious	Low	NI[a]	Low	Low	Critical	Critical
Elm et al [22] (2019)	Low	Low	Low	Moderate	Low	Low	Low	Moderate
Gatsios et al [20] (2020)	Low	Low	Low	Low	Low	Low	Low	Low
Habets et al [26] (2020)	Low	Low	Low	Low	Low	Low	Low	Low
Landers and Ellis [34] (2020)	Low	Moderate	Low	Low	Low	Low	Low	Moderate
Motolese et al [25] (2020)	Low	Low	Low	Low	Low	Low	Low	Low
Wu and Cronin-Golomb [27] (2020)	Low	Low	Low	Low	Low	Low	Low	Low
Horin et al [35] (2019)	Low	Low	Low	Low	NI	Low	Serious	Serious
Kuosmanenet al [24] (2020)	Low	Serious	Moderate	Moderate	Low	Low	Low	Serious

[a]NI: no information.

Table 3. Quality appraisal of the studies: revised Cochrane risk of bias tool for randomized trials.

Author (year)	Randomization process	Deviations from intended interventions	Missing outcome data	Outcome measurements	Selection of the reported results	Overall
Ginis et al [31] (2016)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Lakshminarayana et al [30] (2017)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Ellis et al [23] (2019)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk

Features and Usage of the Mobile Apps

The mobile app system configurations used in this review included 5 types of symptom data collection, reminder, or user interaction functions, given in Table 4. Types of symptom data collection included using the sensor of a smartphone or wearable device, task performance, voice recordings, and self-reported surveys. Among 17 studies, 6 studies collected symptoms using a smartphone accelerometer and gyroscope [16,20,21,24,25,28]. Further, 7 studies used wearable devices [20,23,26,27,31,33,35], which included a smartwatch [20,33], a smart insole [20], an actigraph such as a Fitbit [23], and sensors attached to the ankle [31,35], chest [26], or wrist [26]. Task performance was assessed in 9 studies [16,20,21,24,25,29,30,34,35]. Finger tapping was the most common with 5 studies using it [16,21,25,29,30], followed by cognitive function tests using games or memory tests in 4 studies [20,25,29,30]. There were games such as a ball game [24] and a 9-hole peg game [35] for motor symptom measurement. Task performance also included the sit-to-stand test [34] and the timed up-and-go test [25,34]. Voice data were

XSL•FO
RenderX

collected using the microphone of a smartphone in 2 studies [20,21], and 1 study collected voice data using a head-mounted condenser microphone [35]. Another method of collecting data on symptoms was a self-reported survey [20,22,24,26,27,30,34]. Structured survey tools for electronic patient-reported outcomes [22] and ecological momentary assessments (EMAs) [26,27] were developed. EMAs collect subjective experiences at multiple semirandomized moments during the day to better capture symptom changes.

Functions other than symptom collection were reminders [22,24,30,32,33] or user interactions [23,28,31,34]. Reminder functions, such as symptomatic alerts and medication reminders, were the most common features to assist people with PD in self-care. The user interaction functions provided feedback based on patient activity [23,31,34] or communication with a medical care facility server [28].

The measured outcomes of mobile app usage were participants' satisfaction with the mobile app [25,29,32,34], compliance with using the app [20,22,23,25,27,29,31,34,35], and correlations between the collected symptom data and symptom severity for people with PD [16,20-22,24,26-29,33] (Table 4). Satisfaction with the mobile app was investigated using structured items in various studies. The overall satisfaction rate was 83% to 89.5% [25,29,34], and 1 study reported a rate of 69% [32]. In 1 study, 80% of the users were willing to use the app again because it provided medication reminders via SMS [32], and 97% of the users who used the app to measure motor symptoms responded that the app was useful [29]. Compliance mostly ranged from relatively high (87% to 96%) [20,27,29] to moderate (42.9% to 67.4%) [22,34,35], whereas 1 study reported very low compliance (29.6%) [25]. A study that compared groups with and without a mobile intervention reported no between-group difference in compliance [23]. Several studies reported that the data collected through the app could be used to estimate the severity of motor symptoms [16,21,24,28,29,33].

Patient outcomes were measured in 5 studies. The measured patient outcomes were changes in symptoms or activity levels [23,30,31,34,35], medication adherence [30], and quality of life [23,30,31,34]. Studies have reported an improvement in patient symptoms, activity levels, and gait balance in the mobile app group [31,34]. Further, 2 studies compared activity-level differences between groups using mobile apps and usual interventions; however, there were no differences between the 2 groups in terms of symptoms or activity levels [23,30]. Several studies provided medication reminders using apps, but only 1 study measured medication adherence. This study reported that medication reminders sent using apps led to improved medication adherence [30]. Some studies that measured quality of life reported improvement [23,34], but others did not [30,31].

XSL•FO
RenderX

Table 4. Features and usage of the mobile apps in the included studies.

Study (year)	Features of the mobile app							Outcome measurements			
	Type of symptom data collection					Function		Satisfaction	Feasibility	Symptom severity	Patient outcomes
	Smartphone sensor	Task performance	Voice data	Wearable device	Self-report	Reminder	User interaction				
Keränen and Liikkanen [32] (2013)						✓ SMS		✓			
Pan et al [28] (2015)	✓						✓			✓	
Kassavetis et al [16] (2015)	✓	✓								✓	
Lee et al [29] (2016)		✓ CIT[a]						✓	✓	✓	
Silva de Lima et al [33] (2018)				✓		✓				✓	
Zhan et al [21] (2018)	✓	✓	✓							✓ mPDS[b]	
Elm et al [22] (2019)					✓ ePROs[c]	✓			✓	✓	
Gatsios et al [20] (2020)	✓	✓	✓	✓	✓				✓	✓	
Habets et al [26] (2020)				✓	✓ EMA[d]					✓	
Landers and Ellis [34] (2020)		✓			✓		✓	✓	✓		✓
Motolese et al [25] (2020)	✓	✓ CIT						✓	✓		
Wu and Cronin-Golomb [27] (2020)				✓	✓ EMA				✓	✓	
Horin et al [35] (2019)		✓ Game	✓	✓					✓		✓
Kuosmanen et al [24] (2020)	✓	✓ Game			✓	✓				✓	
Ginis et al [31] (2016)				✓			✓		✓		✓
Lakshminarayana et al [30] (2017)		✓ Game			✓	✓					✓
Ellis et al [23] (2019)				✓			✓		✓		✓

[a]CIT: cognitive interference test.

[b]mPDS: mobile Parkinson disease score.

[c]ePROs: electronic patient-reported outcomes.

[d]EMA: ecological momentary assessment.

Self-care Maintenance

The use of mobile apps for self-care maintenance in this review encompassed medication adherence and physical activity, as indicated in Table 5. Among the 17 studies, 6 were related to medication [22,24,26,30,32,33]. These included 1 RCT [30], 1 quasi-experimental study [24], and 4 observational studies. Of these, 5 studies provided medication reminders via SMS [32] or web push notifications in the apps [22,24,30,33] to promote medication adherence according to a preset medication time. Studies using web push notifications also recorded medication tracking through responses to medication reminders. Another

XSL•FO
RenderX

study collected data on medication intake through EMAs [26]. As outcome measures, studies evaluated medication adherence using self-report questionnaires, participants' satisfaction, as well the relationship between symptom fluctuations or severity and medication intake [24,26,30,32,33]. Another study provided notifications to promote medication adherence through a mobile app, but it did not measure the relevant outcomes [22].

Physical activity was measured in 3 studies among which 2 studies provided tailored exercises to each participant through a mobile app [23,34], and another study consisted of an exercise program for 30 minutes to improve gait, speech, and dexterity symptoms [35]. There was an observational study [34], a quasi-experimental study [35], and an RCT [23]. Landers and Ellis [34] provided tailored video-guided exercises using a proprietary algorithm based on motor symptom data collected through the app. Ellis et al [23] compared the delivery of a prescribed set of exercises with and without mHealth technology. All studies collected information on motor symptoms to measure symptom- and activity-level changes and evaluated the feasibility of the mobile apps based on compliance. Patient outcomes such as quality of life were evaluated in 2 studies [23,34].

XSL•FO
RenderX

Table 5. Self-management characteristics of the mobile apps.

Authors (year)	Self-care maintenance	Self-care monitoring											Self-care management
		Motor symptoms						Nonmotor symptoms					
	PA[a]	TA[b]	Tr.[c]	Rig.[d]	BK[e]	PI[f]	Others	SA[g]	NS[h]	SD[i]	AD[j]	Others	
Keränen and Liikkanen [32] (2013)		✓											
Pan et al [28] (2015)			✓			✓							
Kassavetis et al [16] (2015)			✓		✓								
Lee et al [29] (2016)			✓		✓				✓ Cognition				
Silva de Lima et al [33] (2018)		✓					✓ PA level						
Zhan et al [21] (2018)					✓	✓ Gait	✓ Speech						
Elm et al [22] (2019)		✓	✓	✓	✓	✓ Gait	✓ Speech				✓ Constipation		
Gatsios et al [20] (2020)			✓		✓	✓	✓ Speech/PA level		✓ Cognition/emotion	✓			
Habets et al [26] (2020)		✓	✓	✓	✓	✓	✓ Speech/PA level	✓ Hallucinations	✓ Emotion	✓		✓ Fatigue	
Landers and Ellis [34] (2020)	✓		✓			✓	✓ Fall/PA level						
Motolese et al [25] (2020)			✓		✓	✓ Gait			✓ Cognition				
Wu and Cronin-Golomb [27] (2020)									✓ Cognition/emotion	✓		✓ Fatigue	
Horin et al [35] (2019)	✓ Gait, speech, dexterity		✓		✓	✓ Gait	✓ Speech						
Kuosmanen et al [24] (2020)		✓	✓				✓ Dyskinesia						
Ginis et al [31] (2016)						✓ Gait							✓ Gait training
Lakshminarayana et al [30] (2017)		✓			✓		✓ PA level		✓ Cognition/emotion	✓		✓ Pain	
Ellis et al [23] (2019)	✓						✓ PA level						

[a]PA: physical activity.

XSL•FO
RenderX

[b]TA: treatment adherence.
[c]Tr.: tremor.
[d]Rig.: rigidity.
[e]BK: bradykinesia.
[f]PI: postural instability.
[g]SA: sensory abnormalities.
[h]NS: neuropsychiatric symptoms.
[i]SD: sleep disorder.
[j]AD: autonomic dysfunction.

Self-care Monitoring

Symptoms were monitored in 16 studies. Among them, 7 studies involved self-care monitoring (ie, without self-care maintenance or self-care management) (Table 5). Self-care monitoring assessed the motor and nonmotor symptoms of PD. The most frequently monitored motor symptom was tremor [16,20,22,24-26,28,29,34,35], followed by bradykinesia [16,20-22,25,26,29,30,35], and postural instability and gait [20-22,25,26,28,31,34,35]. Data on rigidity were collected in 2 studies [22,26]. In addition to typical motor symptoms, speech [20-22,26,35], physical activity [20,23,26,30,33,34], and dyskinesia [24] were monitored. Although not technically a motor symptom, fall events [34] were also monitored. Different methods were used to monitor each motor symptom. Smartphones or wearable accelerometers and gyroscopes were mainly used to collect data on tremor [16,20,24-26,28,35], postural instability, and gait symptoms [20,21,28,31,35]. Bradykinesia was usually assessed using task performance such as finger tapping on the screen [16,21,25,29,30], or a 9-hole peg game designed to arouse the patients' interest [35]. Postural instability and tremor were also monitored through performance tasks. Postural instability was assessed by having participants perform the sit-to-stand test [34] and the timed up-and-go test [25,34]. Tremor data were collected using a ball game [24] or rapid alternating movements of the hand holding a smartphone [29]. Rigidity was assessed using self-reported questionnaires only [22,26]. Symptoms related to speech were assessed by self-reports on the severity of symptoms [22,26] or by collecting voice data using a smartphone's microphone or a head-mounted condenser microphone and a digital recorder [20,21,35]. Fall event and dyskinesia data were collected through self-reports. The physical activity level was assessed using self-report questionnaires [26,30,34] or wearable devices [20,23,33].

Among the 7 studies involving self-care monitoring of nonmotor symptoms, neuropsychiatric symptoms (eg, those related to cognition or emotion) were the most common, appearing in 6 studies [20,25-27,29,30]. Symptoms related to sleep disorders were tracked in 4 studies [20,26,27,30]. Other studies gathered information on fatigue [26,27], constipation [22], hallucinations [26], and pain [30]. All nonmotor symptom data were collected using self-reporting questionnaires, except for data on sleep symptoms and cognitive symptoms, which were investigated objectively using wearable devices and task performance, respectively. Sleep data, such as sleep duration and wakefulness, were automatically collected through wearable devices, such as actigraphs [27] or smart watches [20]. Cognition data were collected using task performance, such as cognitive interference tests, memory tests, and cognitive games [20,25,29,30].

Outcomes in self-care monitoring included motor symptom severity estimation from the mobile app data. The severity of symptoms was evaluated in comparison with the clinical scales used in PD such as the Unified Parkinson's Disease Rating Scale. Tremor was the most frequently assessed symptom [16,20,24,28,29], followed by bradykinesia [16,21,29]. The mobile Parkinson disease score and ePROs were developed to measure motor symptoms through the mobile apps [21,22]. The results were compared with clinical data such as the Unified Parkinson's Disease Rating Scale.

Self-care Management

There was 1 study related to self-care management that conducted an RCT with a gait symptom improvement program [31]. The study participants performed walking exercises at least 3 times a week for 30 minutes according to the researchers' instructions. The intervention group members were additionally provided audio biofeedback to improve their balance, gait speed, stride length, and cadence based on the symptoms collected through the sensors on their ankles. This study assessed endurance and quality of life to compare the effectiveness of the gait improvement program with that of conventional gait training.

Discussion

Principal Findings

This review aimed to explore the types, characteristics, and outcomes of mobile apps for self-care in people with PD. Even though mHealth apps have been used widely and positive awareness has grown in the past several years [36], only 17 studies were confirmed as novel studies in the present review. This suggests that the usage of mobile apps for self-care by people with PD is in the early stage. Most studies were observational, whereas a few studies investigated the effects of mobile apps on self-care. There were 3 RCTs, which are insufficient to evaluate the effectiveness of mobile apps used for self-care in people with PD. Most studies investigated self-care monitoring, followed by self-care maintenance and self-care management. These results suggest that the usage of mobile apps for self-care in people with PD is focused on self-care monitoring. Self-care monitoring is important to provide a direction for self-care maintenance and management behaviors in people with PD [6]. Self-care refers to self-monitoring of symptom changes and a series of processes for maintaining a healthy life. Self-care monitoring must be

XSL•FO
RenderX

accompanied by health-promoting behaviors and responses to changes in symptoms [6]. However, almost half of the studies focused only on self-care monitoring [16,20,21,25,27-29].

Features and Usage of the Mobile Apps

Self-care mobile apps for people with other chronic illnesses focused on medication reminders, patient-provider communication, data collection, and transfers of patient outcomes [37]. Specialized software programs or applications were used to check symptoms, connect with patients and diabetes educators in real time, or record a food diary; studies have also deployed wireless or Bluetooth-compatible devices to transfer data automatically from blood pressure monitors, blood glucose meters, electrocardiograms, and scales [37]. Mobile apps for PD use specialized software or applications to generate medication reminders, track symptom data, and facilitate communication between patients and medical care facility servers. However, the most notable mobile apps for people with PD involve using the sensors of smartphones or wearable devices. Accelerometers and gyroscopes of smartphones or wearable devices have advanced from a technological standpoint in that they can effectively capture tremors, postural instability changes, and minute differences in the positions of people with PD [20,28,35]. Studies have used smartwatches or actigraphy to automatically collect sleep data in people with PD [20,27]. According to a qualitative study examining users' perceptions of mHealth apps, many participants preferred tracking technologies based on sensors, such as accelerometers and gyroscopes [36]. Data collection based on sensors or task performance can partially solve the problem of unreliable self-reported data in tracing. Compared to the sensors of smartphones or wearables that would automatically collect data, performance tasks or self-reported questionnaires require the patient to input information directly. Manually inputting data takes time and effort, which could decrease compliance with app usage. However, some symptoms can be monitored only through performance tasks or self-reporting.

Most studies in this review measured adherence to mobile apps, which can be linked to clinical symptom assessment in people with PD. Compliance is an important technology-related issue for interventions using mobile apps. The study with the lowest compliance reported that participants dropped out due to difficulties using smartphones, clinical symptoms, or lack of time [25]. Digital literacy was a factor associated with the use of mobile apps [38]. People with a lower socioeconomic status and those who were older had low awareness of health apps or faced difficulties in using them [36]. A study in this review reported that motor-related aspects of daily living, patients' self-rated health status, and caregivers' burden were the determinants of compliance [20]. These factors could be barriers hindering continued app usage. Elm et al [22] reported declining amounts of streaming and reporting over time, specifically after the first 3 months. As a study pointed out, patients preferred straightforward and simple methods [36]. People with PD might experience difficulties using a smartphone because they are older and have motor symptoms. User-centered interface configurations, which consider the characteristics such as the age and disease of the users, should be considered to increase compliance.

PD involves various motor symptoms due to a marked decrease in the neurotransmitter dopamine, which needs accurate assessment of disease-related symptoms [4]. The studies included in this review showed that data collected through mobile apps could effectively assess disease severity in people with PD. This finding suggests the possibility of regular home-based assessments to capture symptom changes between follow-up visits with clinicians.

The goal of self-care in chronic illness is to maintain optimal living with the disease, which means maintaining one's health status, improving well-being and quality of life, reducing health care use, and decreasing mortality and symptom burden [6]. It is necessary to assess the clinical outcomes related to self-care to evaluate the effects of using mobile apps for self-care. In this regard, 3 systematic reviews about self-care apps for people with chronic illnesses (ie, chronic lung disease, cardiovascular disease, and diabetes mellitus) identified effectiveness in terms of clinical outcomes such as changes in physical function and clinical results (eg, 6-minute walking test, hemoglobin A_{1c}, blood pressure, blood glucose, or body weight), compliance with a treatment regimen, performance of self-care tasks, and quality of life [11,37,38]. Among the studies considered in this review, 5 assessed clinical outcomes related to self-care. The results of these studies showed that the usage of mobile apps in patients with PD was still insufficient to confirm whether patient outcomes such as changes in symptoms or activity levels, medication adherence, and quality of life had improved.

Self-care Maintenance

It is known that the motor symptoms of PD can be effectively controlled by medications [4]; therefore, medication adherence is very important in PD. It is not surprising that the first study on mobile apps for self-care in PD involved medication reminders to promote medication adherence [32]. Web push notifications are effective in tracking medication adherence, whereas SMS can only provide medication reminders. Recording responses to medication reminders is a more objective method for assessing medication adherence than a self-reporting questionnaire. However, no studies analyzed collected medication records to assess medication adherence. This finding suggests that future research needs to focus on symptom changes according to medication adherence rather than subjectively measuring adherence.

Physical activity has been established as the most effective way of improving physical and cognitive functions in people with PD [39]. Many PD patients struggle to participate in exercise programs due to their functional limitations and abilities [34]. They may sometimes be motivated to perform healthy behaviors but may not know the right way to perform them [36]. Many people using health-promoting apps value personalized and tailored information [36]. People with PD need personalized coaching and specific exercise planning programs tailored to their functional abilities. A study found that a customized exercise program using a mobile app could be safely and effectively provided to people with PD who could not regularly participate in exercise programs due to symptoms or functional

XSL•FO
RenderX

changes [34]. Various face-to-face interventions focused on improving fatigue, stress, sleep, and nutrition were provided to maintain a healthy lifestyle via self-management [7,8]. However, the interventions using mobile apps focused mostly on medication adherence and physical activity.

Self-care Monitoring

Among the motor symptoms, tremor, bradykinesia, postural instability, and gait were monitored frequently. The results show that monitoring these symptoms has important implications for the management of PD. Rigidity, which is referred to as a major motor symptom in the literature, was assessed less frequently than other symptoms [4]. A reason for this might be that rigidity can only be measured through self-reporting, unlike symptoms such as tremor, bradykinesia, postural instability, and gait, which can be objectively measured through wearable devices or task performance.

People with PD experience various nonmotor symptoms in addition to motor symptoms [3,4]. Similar to motor symptoms, nonmotor symptoms contribute toward deteriorating quality of life [5]. This review found that self-care monitoring using a mobile app in people with PD often focused more on monitoring motor symptoms than nonmotor symptoms. The nonmotor symptoms experienced by people with PD include cognitive impairment, sleep problems, urinary problems, pain, fatigue, and constipation [5]. This review showed that cognitive or emotional impairment and sleep were the main nonmotor symptoms monitored using mobile apps. Except for cognitive impairment and sleep disturbance, other nonmotor symptoms are subjective and difficult to assess. As nonmotor symptoms have a significant impact on the quality of life of patients with PD, they should be monitored using various structured tools.

Self-care Management

Previous studies reported interventions applied for self-care management in people with diabetes mellitus or hypertension, such as goal management, motivational feedback, and health coaching through mobile apps [11,12]. These interventions have been confirmed in face-to-face interventions for self-care management. Only 1 study analyzed a self-care management intervention through a mobile app for people with PD. The study involved gait training with audio biofeedback [31]. Because this app provided feedback according to the individual's gait performance, it had a corrective effect on gait symptoms. Self-care management interventions function as a navigator to change health practices or seek medical resources in a timely manner when the symptoms occur. This review confirms that self-care management interventions using mobile apps in people with PD are highly insufficient. There is a need to develop mobile apps for patients with PD that can guide medication adherence, physical activity enhancement, or use of health care resources when symptom changes occur.

Strengths and Limitations

Several reviews on mobile apps for people with PD have been conducted. However, previous reviews compared the iOS and Android operating systems or analyzed the potential usability of these apps for assessing and treating PD [14,15]. In contrast, we focused on analyzing the usage of mobile apps for self-care. As PD is a progressive disease, self-care is very important for maintenance, monitoring, and symptom management. This review makes a meaningful contribution to existing research by identifying the strengths and weaknesses related to the usage and development of mobile apps for self-care in people with PD. Nevertheless, several limitations should be noted. First, owing to the low number of RCTs, we could not compare the effectiveness of mobile apps for self-care. Second, because we excluded protocols, studies limited to only technical issues, and articles published in non-English languages, there was a potential bias in literature selection that could have influenced the interpretation of the results.

Implications

We found that the motor and nonmotor symptoms of patients with PD could be continuously monitored through mobile apps and that disease severity could be estimated using the collected data. Smartphone sensors and wearable devices measured motor symptoms objectively. A structured tool could be a possible option to collect nonmotor symptom data. Studies on mobile apps for patients with PD showed that interventions targeting medication adherence or physical activity were applicable. There is a need to develop self-care interventions that organically connect health promotion behaviors, symptom monitoring, and behavior changes with the usage of mobile apps in patients with PD.

Conclusions

This review identified that the usage of mobile apps for self-care in people with PD focused only on disease-specific characteristics and did not involve approaches to symptom management. These results imply that future research on mobile app development for people with PD should involve strategies for self-care management and maintenance based on symptom monitoring. Further research is needed to build evidence to support the usage of mobile apps for self-care in people with PD and evaluate the effects of such apps on quality of life and symptom improvement.

Acknowledgments

This study was supported by the Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Education (grants 2018R1D1A1A02085559 and 2020R1A6A1A03041989), the Brain Korea 21 FOUR Project funded by the NRF, and a 2021 International Joint Research Grant from Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing. Authors SY and EK received a scholarship from the Brain Korea 21 FOUR Project funded by the NRF, Yonsei University College of Nursing.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Searching terms.
[DOCX File , 23 KB - mhealth_v10i1e33944_app1.docx]

References

1. GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):459-480 [FREE Full text] [doi: 10.1016/S1474-4422(18)30499-X] [Medline: 30879893]
2. Pringsheim T, Jette N, Frolkis A, Steeves TD. The prevalence of Parkinson's disease: a systematic review and meta-analysis. Mov Disord 2014 Nov;29(13):1583-1590. [doi: 10.1002/mds.25945] [Medline: 24976103]
3. Schapira AH, Chaudhuri KR, Jenner P. Non-motor features of Parkinson disease. Nat Rev Neurosci 2017 Jul;18(7):435-450. [doi: 10.1038/nrn.2017.62] [Medline: 28592904]
4. Kalia LV, Lang AE. Parkinson's disease. Lancet 2015 Aug;386(9996):896-912. [doi: 10.1016/S0140-6736(14)61393-3] [Medline: 25904081]
5. Kadastik-Eerme L, Rosenthal M, Paju T, Muldmaa M, Taba P. Health-related quality of life in Parkinson's disease: a cross-sectional study focusing on non-motor symptoms. Health Qual Life Outcomes 2015 Jun;13:83 [FREE Full text] [doi: 10.1186/s12955-015-0281-x] [Medline: 26088201]
6. Riegel B, Jaarsma T, Lee CS, Strömberg A. Integrating symptoms into the middle-range theory of self-care of chronic illness. ANS Adv Nurs Sci 2019 Jul;42(3):206-215 [FREE Full text] [doi: 10.1097/ANS.0000000000000237] [Medline: 30475237]
7. Kessler D, Liddy C. Self-management support programs for persons with Parkinson's disease: an integrative review. Patient Educ Couns 2017 Oct;100(10):1787-1795. [doi: 10.1016/j.pec.2017.04.011] [Medline: 28465112]
8. Owen CL, Ibrahim K, Dennison L, Roberts HC. Falls self-management interventions for people with Parkinson's disease: a systematic review. J Parkinsons Dis 2019 Feb;9(2):283-299 [FREE Full text] [doi: 10.3233/JPD-181524] [Medline: 30958315]
9. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013 May;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
10. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013 Jan;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
11. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. PLoS One 2016 Nov;11(11):e0166718 [FREE Full text] [doi: 10.1371/journal.pone.0166718] [Medline: 27861583]
12. Liu K, Xie Z, Or CK. Effectiveness of mobile app-assisted self-care interventions for improving patient outcomes in type 2 diabetes and/or hypertension: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Aug;8(8):e15779 [FREE Full text] [doi: 10.2196/15779] [Medline: 32459654]
13. Denis F, Lethrosne C, Pourel N, Molinier O, Pointreau Y, Domont J, et al. Randomized trial comparing a web-mediated follow-up with routine surveillance in lung cancer patients. J Natl Cancer Inst 2017 Dec;109(9):1-8. [doi: 10.1093/jnci/djx029] [Medline: 28423407]
14. Estévez S, Cambronero M, García-Ruiz Y, Llana DL. Mobile applications for people with Parkinson's disease: a systematic search in app stores and content review. J Univers Comput Sci 2019 Jul;25(7):740-763. [doi: 10.3217/jucs-025-07-0740]
15. Linares-Del Rey M, Vela-Desojo L, Cano-de la Cuerda R. Mobile phone applications in Parkinson's disease: a systematic review. Neurologia 2019 Jan;34(1):38-54 [FREE Full text] [doi: 10.1016/j.nrl.2017.03.006] [Medline: 28549757]
16. Kassavetis P, Saifee TA, Roussos G, Drougkas L, Kojovic M, Rothwell JC, et al. Developing a tool for remote digital assessment of Parkinson's disease. Mov Disord Clin Pract 2016 Jan;3(1):59-64 [FREE Full text] [doi: 10.1002/mdc3.12239] [Medline: 30363542]
17. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
18. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
19. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]

XSL•FO
RenderX

20. Gatsios D, Antonini A, Gentile G, Marcante A, Pellicano C, Macchiusi L, et al. Feasibility and utility of mHealth for the remote monitoring of Parkinson disease: ancillary study of the PD_manager randomized controlled trial. JMIR Mhealth Uhealth 2020 Jun;8(6):e16414 [FREE Full text] [doi: 10.2196/16414] [Medline: 32442154]
21. Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity. JAMA Neurol 2018 Jul 01;75(7):876-880. [doi: 10.1001/jamaneurol.2018.0809]
22. Elm JJ, Daeschler M, Bataille L, Schneider R, Amara A, Espay AJ, et al. Feasibility and utility of a clinician dashboard from wearable and mobile application Parkinson's disease data. NPJ Digit Med 2019 Sep;2:95 [FREE Full text] [doi: 10.1038/s41746-019-0169-y] [Medline: 31583283]
23. Ellis T, Cavanaugh J, DeAngelis T, Hendron K, Thomas C, Saint-Hilaire M, et al. Comparative effectiveness of mHealth-supported exercise compared with exercise alone for people with Parkinson disease: randomized controlled pilot study. Phys Ther 2019 Feb;99(2):203-216. [doi: 10.1093/ptj/pzy131] [Medline: 30715489]
24. Kuosmanen E, Wolling F, Vega J, Kan V, Nishiyama Y, Harper S, et al. Smartphone-based monitoring of Parkinson disease: quasi-experimental study to quantify hand tremor severity and medication effectiveness. JMIR Mhealth Uhealth 2020 Nov;8(11):e21543. [doi: 10.2196/21543] [Medline: 33242017]
25. Motolese F, Magliozzi A, Puttini F, Rossi M, Capone F, Karlinski K, et al. Parkinson's disease remote patient monitoring during the COVID-19 lockdown. Front Neurol 2020 Oct;11:567413 [FREE Full text] [doi: 10.3389/fneur.2020.567413] [Medline: 33117262]
26. Habets J, Heijmans M, Herff C, Simons C, Leentjens AF, Temel Y, et al. Mobile health daily life monitoring for Parkinson disease: development and validation of ecological momentary assessments. JMIR Mhealth Uhealth 2020 May;8(5):e15628 [FREE Full text] [doi: 10.2196/15628] [Medline: 32339999]
27. Wu JQ, Cronin-Golomb A. Temporal associations between sleep and daytime functioning in Parkinson's disease: a smartphone-based ecological momentary assessment. Behav Sleep Med 2020 Jun;18(4):560-569. [doi: 10.1080/15402002.2019.1629445] [Medline: 31199163]
28. Pan D, Dhall R, Lieberman A, Petitti DB. A mobile cloud-based Parkinson's disease assessment system for home-based monitoring. JMIR Mhealth Uhealth 2015 Mar;3(1):e29 [FREE Full text] [doi: 10.2196/mhealth.3956] [Medline: 25830687]
29. Lee W, Evans A, Williams DR. Validation of a smartphone application measuring motor function in Parkinson's disease. J Parkinsons Dis 2016 Apr;6(2):371-382. [doi: 10.3233/JPD-150708] [Medline: 27061062]
30. Lakshminarayana R, Wang D, Burn D, Chaudhuri KR, Galtrey C, Guzman NV, et al. Erratum: Author correction: Using a smartphone-based self-management platform to support medication adherence and clinical consultation in Parkinson's disease. NPJ Parkinsons Dis 2017 Nov;3:32 [FREE Full text] [doi: 10.1038/s41531-017-0034-0] [Medline: 29152558]
31. Ginis P, Nieuwboer A, Dorfman M, Ferrari A, Gazit E, Canning CG, et al. Feasibility and effects of home-based smartphone-delivered automated feedback training for gait in people with Parkinson's disease: A pilot randomized controlled trial. Parkinsonism Relat Disord 2016 Jan;22:28-34. [doi: 10.1016/j.parkreldis.2015.11.004] [Medline: 26777408]
32. Keränen T, Liikkanen S. Medication reminder service for mobile phones: an open feasibility study in patients with Parkinson's disease. Telemed J E Health 2013 Nov;19(11):888-890. [doi: 10.1089/tmj.2013.0014] [Medline: 24050614]
33. Silva de Lima AL, Evers LJ, Hahn T, de Vries NM, Daeschler M, Boroojerdi B, et al. Impact of motor fluctuations on real-life gait in Parkinson's patients. Gait Posture 2018 May;62:388-394. [doi: 10.1016/j.gaitpost.2018.03.045] [Medline: 29627498]
34. Landers MR, Ellis TD. A mobile app specifically designed to facilitate exercise in Parkinson disease: single-cohort pilot study on feasibility, safety, and signal of efficacy. JMIR Mhealth Uhealth 2020 Oct;8(10):e18985 [FREE Full text] [doi: 10.2196/18985] [Medline: 33016887]
35. Horin AP, McNeely ME, Harrison EC, Myers PS, Sutter EN, Rawson KS, et al. Usability of a daily mHealth application designed to address mobility, speech and dexterity in Parkinson's disease. Neurodegener Dis Manag 2019 Apr;9(2):97-105 [FREE Full text] [doi: 10.2217/nmt-2018-0036] [Medline: 30998100]
36. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
37. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
38. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
39. Lauzé M, Daneault J, Duval C. The effects of physical activity in Parkinson's disease: a review. J Parkinsons Dis 2016 Oct;6(4):685-698. [doi: 10.3233/JPD-160790] [Medline: 27567884]

Abbreviations

EMAs: ecological momentary assessments
mHealth: mobile health

XSL•FO
RenderX

PD: Parkinson disease
RCTs: randomized controlled trials
ROBINS-I: risk of bias in non-randomized studies of interventions
RoB2: revised Cochrane risk of bias tool for randomized trials

Edited by G Eysenbach; submitted 30.09.21; peer-reviewed by L Hess; comments to author 06.11.21; revised version received 19.11.21; accepted 30.11.21; published 21.01.22.

Please cite as:
Lee J, Yeom I, Chung ML, Kim Y, Yoo S, Kim E
Use of Mobile Apps for Self-care in People With Parkinson Disease: Systematic Review
JMIR Mhealth Uhealth 2022;10(1):e33944
URL: https://mhealth.jmir.org/2022/1/e33944
doi:10.2196/33944
PMID:35060910

©JuHee Lee, Insun Yeom, Misook L Chung, Yielin Kim, Subin Yoo, Eunyoung Kim. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 21.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

The Impact of Wearable Technologies in Health Research: Scoping Review

Sophie Huhn[1]; Miriam Axt[1], BA; Hanns-Christian Gunga[2], Prof Dr; Martina Anna Maggioni[2,3], PD, PhD; Stephen Munga[4], PhD; David Obor[4], MSc; Ali Sié[1,5], PhD; Valentin Boudo[5], MSc; Aditi Bunker[1], Dr sc hum; Rainer Sauerborn[1], Prof Dr; Till Bärnighausen[1,6,7], Prof Dr; Sandra Barteit[1], Dr sc hum

[1]Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg University, Heidelberg, Germany

[2]Charité - Universitätsmedizin Berlin, Institute of Physiology, Center for Space Medicine and Extreme Environment, Berlin, Germany

[3]Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milano, Italy

[4]Kenya Medical Research Institute, Kisumu, Kenya

[5]Centre de Recherche en Santé Nouna, Nouna, Burkina Faso

[6]Harvard Center for Population and Development Studies, Cambridge, MA, United States

[7]Africa Health Research Institute, KwaZulu-Natal, South Africa

Corresponding Author:
Sophie Huhn
Heidelberg Institute of Global Health
Heidelberg University Hospital
Heidelberg University
Im Neuenheimer Feld 130 3
Heidelberg, 69120
Germany
Phone: 49 6221 5634030
Email: sophie.huhn@uni-heidelberg.de

Abstract

Background: Wearable devices hold great promise, particularly for data generation for cutting-edge health research, and their demand has risen substantially in recent years. However, there is a shortage of aggregated insights into how wearables have been used in health research.

Objective: In this review, we aim to broadly overview and categorize the current research conducted with affordable wearable devices for health research.

Methods: We performed a scoping review to understand the use of affordable, consumer-grade wearables for health research from a population health perspective using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) framework. A total of 7499 articles were found in 4 medical databases (PubMed, Ovid, Web of Science, and CINAHL). Studies were eligible if they used noninvasive wearables: worn on the wrist, arm, hip, and chest; measured vital signs; and analyzed the collected data quantitatively. We excluded studies that did not use wearables for outcome assessment and prototype studies, devices that cost >€500 (US $570), or obtrusive smart clothing.

Results: We included 179 studies using 189 wearable devices covering 10,835,733 participants. Most studies were observational (128/179, 71.5%), conducted in 2020 (56/179, 31.3%) and in North America (94/179, 52.5%), and 93% (10,104,217/10,835,733) of the participants were part of global health studies. The most popular wearables were fitness trackers (86/189, 45.5%) and accelerometer wearables, which primarily measure movement (49/189, 25.9%). Typical measurements included steps (95/179, 53.1%), heart rate (HR; 55/179, 30.7%), and sleep duration (51/179, 28.5%). Other devices measured blood pressure (3/179, 1.7%), skin temperature (3/179, 1.7%), oximetry (3/179, 1.7%), or respiratory rate (2/179, 1.1%). The wearables were mostly worn on the wrist (138/189, 73%) and cost <€200 (US $228; 120/189, 63.5%). The aims and approaches of all 179 studies revealed six prominent uses for wearables, comprising correlations—wearable and other physiological data (40/179, 22.3%), method evaluations (with subgroups; 40/179, 22.3%), population-based research (31/179, 17.3%), experimental outcome assessment (30/179, 16.8%), prognostic forecasting (28/179, 15.6%), and explorative analysis of big data sets (10/179, 5.6%). The most frequent strengths of affordable wearables were validation, accuracy, and clinical certification (104/179, 58.1%).

XSL•FO
RenderX

Conclusions: Wearables showed an increasingly diverse field of application such as COVID-19 prediction, fertility tracking, heat-related illness, drug effects, and psychological interventions; they also included underrepresented populations, such as individuals with rare diseases. There is a lack of research on wearable devices in low-resource contexts. Fueled by the COVID-19 pandemic, we see a shift toward more large-sized, web-based studies where wearables increased insights into the developing pandemic, including forecasting models and the effects of the pandemic. Some studies have indicated that big data extracted from wearables may potentially transform the understanding of population health dynamics and the ability to forecast health trends.

(JMIR Mhealth Uhealth 2022;10(1):e34384) doi:10.2196/34384

KEYWORDS

wearable; consumer-grade wearables; commercially available wearables; public health; global health; population health; fitness trackers; big data; low-resource setting; tracker; review; mHealth; research; mobile phone

Introduction

Background

Wearable devices hold great promise, particularly for data generation for cutting-edge health research, and their demand has risen considerably in the last few years [1-3].

Noninvasive, consumer-grade wearables (hereafter *wearables*) may provide manifold advantages for health research; they are generally unobtrusive, less expensive than gold standard research devices [4], comfortable to wear [5], and affordable for consumers [6]. In recent years, the quality and accuracy of wearables have improved [7,8], resulting in more clinically approved certifications [9]. Wearables can measure long-term data in the naturalistic environment of study participants, allowing for ecologic momentary assessments [10,11]. Therefore, wearables are valuable developments, particularly for generating data for health research in large study populations, that is, global health or epidemiological studies, or in low-income contexts [6,9,12].

One example of a large study is the so-called *Datenspende* study by the Robert Koch Institute, the German research institute for disease control and prevention, which aims to tackle the COVID-19 (corona virus disease) pandemic with anonymous data donations acquired through wearables [13]. On the basis of the study by Radin et al [14], researchers used wearable data to calculate the regional probability of COVID-19 outbreaks incorporating data on pulse, physical activity (PA), and sleep, as well as weather data. Using a large sample size exceeding half a million participants, they forecasted the number of COVID-19 infections for the preceding 4 days. The Apple Heart Study [15] is another example that was a breakthrough for showing that wearable devices may detect atrial fibrillation (AF) and foster a discussion of potentials and limitations with regard to health care providers, researchers, and members of the media and economy [16,17].

Apart from these 2 examples, wearables are applied in diverse fields of health, including acoustic, gastrointestinal sensors for ileus prediction [18]; UV sun exposure [19]; heat-related illness measurements [20]**;** electrolyte monitoring, for example, for cystic fibrosis or training management [21,22]; early warning of AF with a wearable ring [23]; generation of electrocardiograms (ECGs) [15]; measurement of cardiopulmonary resuscitation quality [24]; measurement of continuous noninvasive blood glucose [25], as well as smart inhalers and activity trackers for asthma monitoring [26].

Numerous reviews and studies have investigated validation and accuracy, particularly for specific affordable wearables, comparing these to the gold standard measurements [21] or comparing evidence in a meta-analysis [8]. Many studies have focused on novel technologies, presenting prototypes, or investigating the feasibility and acceptance of a wearable device in a specific setting [3,27]. Similarly, reviews on the application and potential of wearables have focused on (1) specific wearable devices or specific wearable measurements, for example, only smartwatches [4] or only sleep measurements [28] or (2) applications of specific medical fields and interventions, for example, only for diagnosis and treatment in cardiological conditions [29] or wearables as an intervention to promote PA in patients with oncologic conditions [30]. Among these publications, we identified a lack of aggregated insight for wearable use in health research and its respective strengths and shortcomings.

Objectives

With this scoping review, we aim to overview and categorize the current research conducted on wearable devices.

Methods

Overview

We conducted a scoping review to explore the applications of affordable wearables worn on wrists, arms, chests, or waists, which constitute the characteristic locations [31]. We focused on the following aspects: (1) demographics; (2) wearable devices and measured vital signs; (3) wearable data and its analysis; (4) reported shortcomings and strengths of wearables; and (5) study aims, results, and types of wearable use. We present our findings in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting standard and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; Multimedia Appendix 1) [32] and the methodological framework of Arksey and O'Malley [33] and Peters et al [34]. A scoping review seemed most appropriate given the broad nature of this subject and the range of potential implementations in the setting of health research.

XSL•FO
RenderX

Eligibility Criteria

We sought to define and characterize the state of affordable wearables for health research. Eligible publications were peer reviewed, published in English, and published after 2013 (after wearables became widely commercially available [1-3]) and had a full-text version available (in instances no full text was available, authors were contacted 3 times with a waiting period of 7 days between each contact before exclusion).

Our review scopes the current information available on affordable, noninvasive wearables, which are (1) worn on the wrist, arm, and chest; (2) measure vital signs; and (3) analyze the generated wearable data for outcome assessment. Validation and qualitative studies were excluded. We focused only on devices that cost <€500 (US $570) per device (1) to allow the affordability of larger studies, for example, where wearable devices need to be provided to study participants via the study and (2) to ensure that wearables are available commercially and (3) intended for consumers. As the definition of vital signs is not distinct [35], we included the following vital signs [9,36,37]: HR, HR variability, ECG measurements or heart rhythm analysis (detection of arrhythmias), blood pressure, blood oxygen, respiratory rate, body temperature, sleep, electrodermal activity, electromyogram measurements, and PA (Textbox 1).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Publications
  - Full text available
  - English language
  - Peer-reviewed articles
  - Published between 2013 and 2020
- Wearable device
  - Commercially available wearable, price <€500 (US $570) per device (Only hardware prices were considered. Software, subscriptions, or similar, which might be necessary for device use, were not included. All prices were captured in the timeframe of this study and therefore are only considered as approximations)
  - Wearables worn on the arm, wrist, chest, and waist
- Outcomes
  - Measuring and analyzing one or more vital sign
  - Range of vital signs as defined in this review, including heart rate, heart rate variability, electrocardiogram measurements or heart rhythm analysis (detection and classification of atrial fibrillation, extrasystoles, and other arrhythmic events), blood pressure, blood oxygen, respiratory rate, body temperature, sleep (time, deepness, etc), electrodermal activity, electromyogram measurements, physical activity (steps, distance covered, intensity, energy expenditure, etc; physical activity included as basic measurements of wearables or very similar or related parameters) [9,36,37].

Exclusion criteria

- Publications
  - Studies not analyzing wearable-generated data for (health) outcome assessment, including studies focusing on (1) accuracy, validation, improvement (algorithms and software); (2) patents; (3) smart clothing; (4) obtrusive wearables (the device comprises obstructive parts or wires, etc); (5) behavior change intervention studies (ie, where the wearable is provided as promotion for more physical activity only and not for health outcome assessment); (6) qualitative studies; or (7) studies with research objectives and outcomes not related to health or a medical condition
- Wearable device
  - Wearable not commercially available (eg, prototype and discontinued)
  - Invasive, obtrusive device (comprising obstructive parts or wires, etc)
  - Prosthesis, smart clothing (sensors in clothing)
- Outcomes
  - Not measuring vital sign, that is, gait, posture, and motion recognition analysis (eg, gesture recognition for sign language)
  - Studies with research objectives and outcomes not related to health or a medical condition

XSL•FO
RenderX

Information Sources and Search

We used PubMed, Ovid, Web of Science, and CINAHL to search peer-reviewed literature using a search string based on the following three concepts: synonyms and medical subject headings terms, including (1) wearables (synonyms, top 15 vendors with most market shares [38-40], or frequently used in research [2,7]), (2) physical wear location of wearables (torso, arm, and wrist), and (3) measurement of vital signs (for full search string see Multimedia Appendix 2 [41]). We manually searched the reference lists for relevant articles.

We imported the identified articles into the literature reference management system Zotero [42] and then into the systematic review management platform Covidence [41]. Literature was screened by 2 independent reviewers. Any disagreements were resolved by discussion between the 2 reviewers (SH and MA) and a third researcher (SB).

Quality Assessment

To assess the quality of the included studies and their various study designs (credibility), we considered the Medical Education Research Study Quality Instrument [43] score as adequate (Multimedia Appendix 3 [14,15,20,44-219]).

Data Synthesis

We conducted data synthesis in accordance with Arksey and O'Malley [33], comprising the analytic framework, analysis of the extent and nature of studies, and thematic analysis. We categorized the findings by title, author, year, country of study, objectives of study, study population, sample size, methods, intervention type, outcomes, and key findings related to the scoping review question [34]. We extracted mutually exclusive groups, including wearable manufacturers, built-in sensors, scope of measurements (vital signs), shortcomings and strengths of wearables mentioned by the authors, the used methods for data analysis, and medical fields.

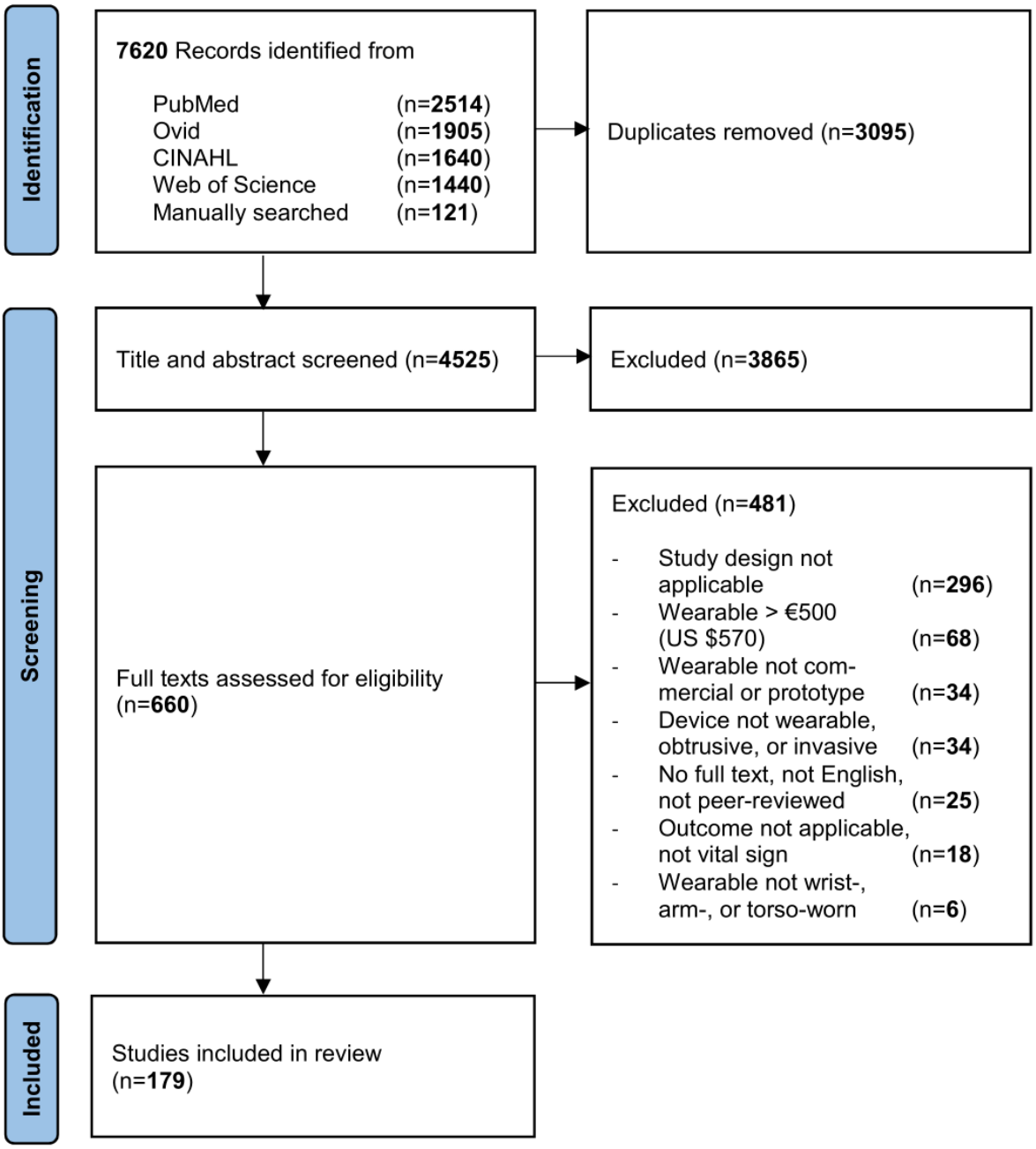
Results

Overview

Our initial search yielded 7499 hits (PubMed: 2514; Ovid: 1905; Web of Science: 1440; CINAHL: 1640) and we identified 121 publications by manual search. Of 7620 total publications, we screened 4525 (59.38%) nonduplicates for title and abstract, leading to the assessment of 660 full-texts. After full-text screening of the 660 articles, we included 179 (27.1%) studies in our review [14,15,20,44-219] (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [220].



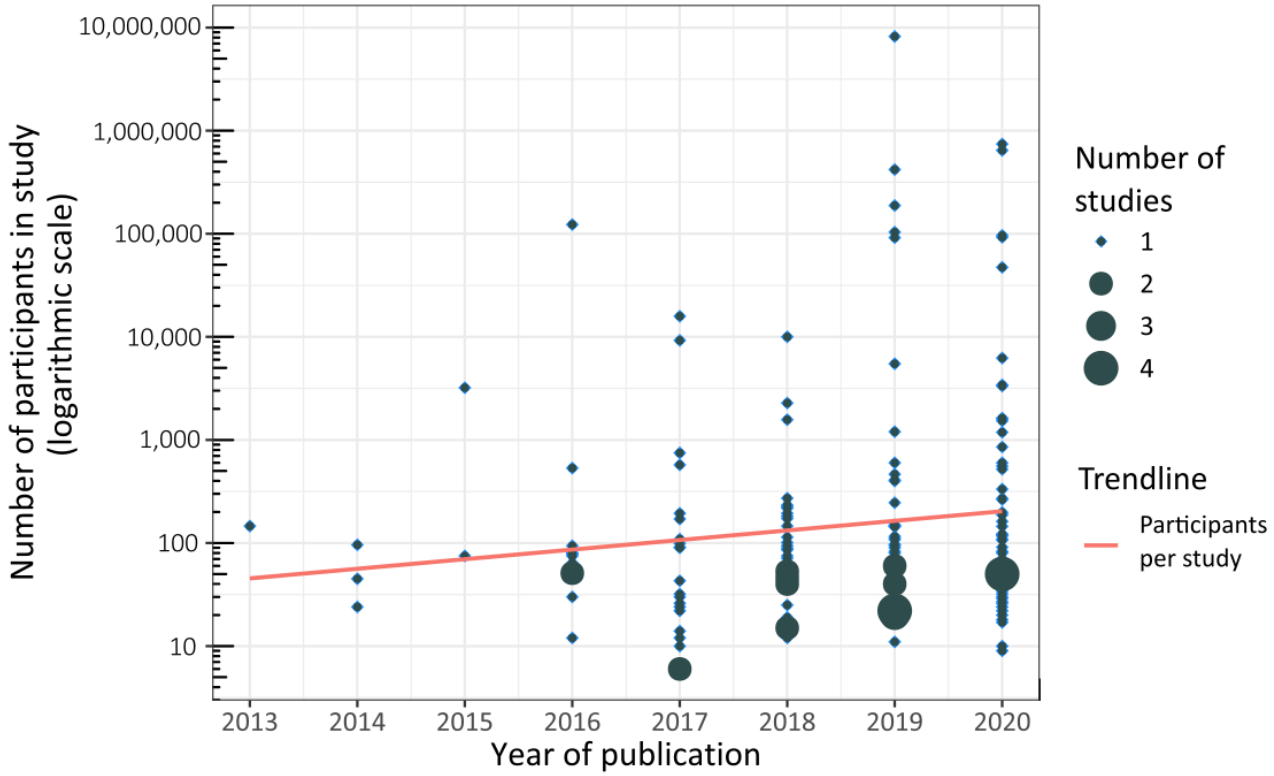
Study Characteristics

Demographics

Between 2013 and 2020, we observed an increase in the number of studies and study participants (Figure 2 and Table 1). The year 2019 featured the largest sample size, and studies were predominantly conducted in North America (Figure 3 [221]). The largest study we identified was conducted in 2019 in North America and included over 8 million participants (75.71%) [153]; the second largest was a European study comprising 742,000 participants (6.85%) [162]. Without the aforementioned, largest study, Europe and Asia would lead in participant numbers and we would see a continuous increase in participant numbers from 2013 to 2020.

XSL•FO
RenderX

Figure 2. Number of studies and study participants (logarithmic scale) per year of study publication. The sizes of the circles visualize the overlapping and number of studies within the year.



XSL•FO
RenderX

Table 1. Characteristics of studies.

Study characteristics	Studies (N=179), n (%)	Participants (N=10,835,733), n, (%)
Year of publication		
2013	1 (0.56)	146 (<0.01)
2014	3 (1.68)	165 (<0.01)
2015	2 (1.12)	3284 (0.03)
2016	14 (7.82)	124,060 (1.14)
2017	21 (11.73)	27,377 (0.25)
2018	34 (18.99)	16,700 (0.15)
2019	48 (26.82)	9,016,909 (83.21)
2020	56 (31.28)	1,647,092 (15.2)
Continents		
North America	94 (52.51)	8,916,888 (82.29)
Europe	50 (27.93)	991,357 (9.15)
Asia	24 (13.41)	925,768 (8.54)
Australia	8 (4.47)	1198 (0.01)
South America	3 (1.68)	522 (<0.01)
Study objectives		
Correlations and influencing factors of study population and outcome data[a]	70 (39.11)	394,296 (3.64)
Population and patient characterization[b]	54 (30.17)	8,315,559 (76.74)
Evaluation of method or intervention	47 (26.26)	2,124,328 (19.6)
Prognostic evaluation[c]	8 (4.5)	1550 (0.01)
Study design		
Cross-sectional study	66 (36.87)	9,780,808 (90.26)
Cohort study	62 (34.64)	628,641 (5.8)
Nonrandomized experimental study	14 (7.82)	724 (0.01)
Randomized controlled trial	11 (6.15)	2332 (0.02)
Method evaluation	8 (4.47)	314,247 (2.9)
Other	7 (3.91)	108,462 (1)
Case control study	7 (3.91)	348 (<0.01)
Mixed methods, feasibility study	4 (2.23)	171 (<0.01)
Medical field of study		
Multidisciplinary and general medicine	43 (24.02)	107,148 (0.99)
Neurology and psychiatry	29 (16.2)	2630 (0.02)
Cardiology, fitness, and sports medicine	28 (15.64)	557,120 (5.14)
Global health, epidemiology, and prevention	19 (10.61)	10,104,217 (93.25)
Gynecology and pediatrics	18 (10.06)	5575 (0.05)
Orthopedics and surgery	16 (8.94)	2749 (0.03)
Pulmonology	13 (7.26)	1326 (0.01)
Other	13 (7.26)	54,968 (0.51)

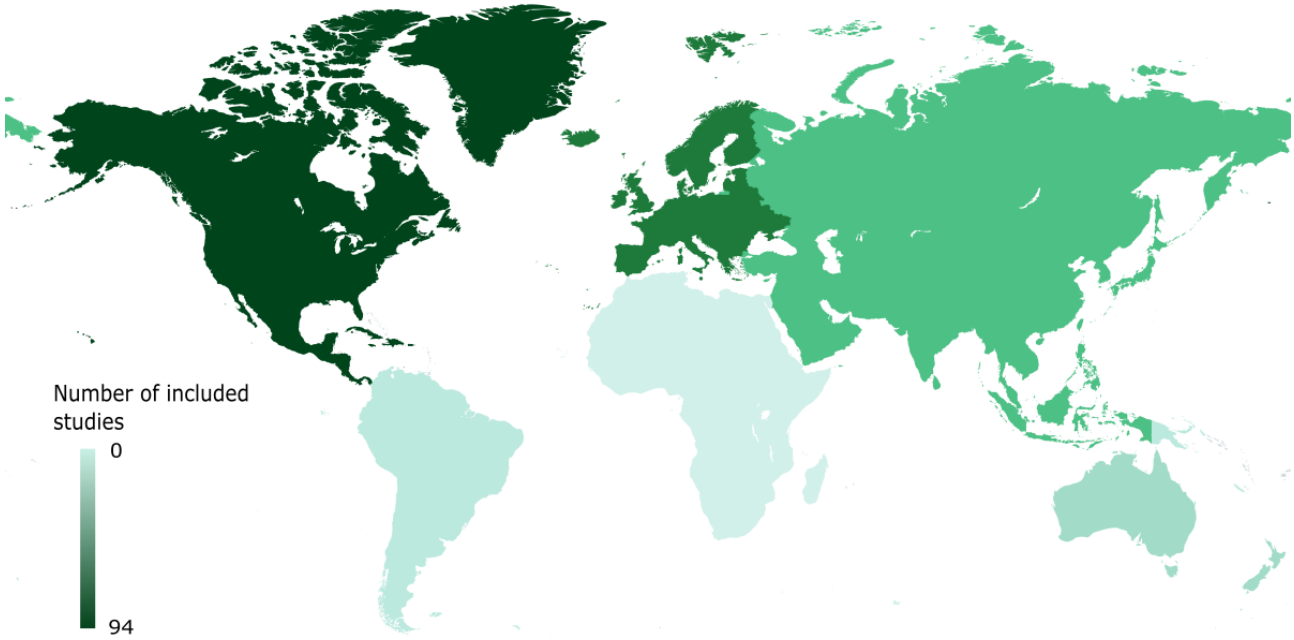
[a]Studies aimed to find associations, correlations, or influencing factors within their study population, study outcomes, and generated data.

[b]Studies aimed to observe and characterize the study population and patients.

[c]Studies aimed to evaluate patient-reported outcomes, health care practices, diagnostics, screenings, and others.

XSL•FO
RenderX

Figure 3. Included studies per continent. The colors of the continents visualize the number of included studies published on the respective continent (created with Mapchart [221]).



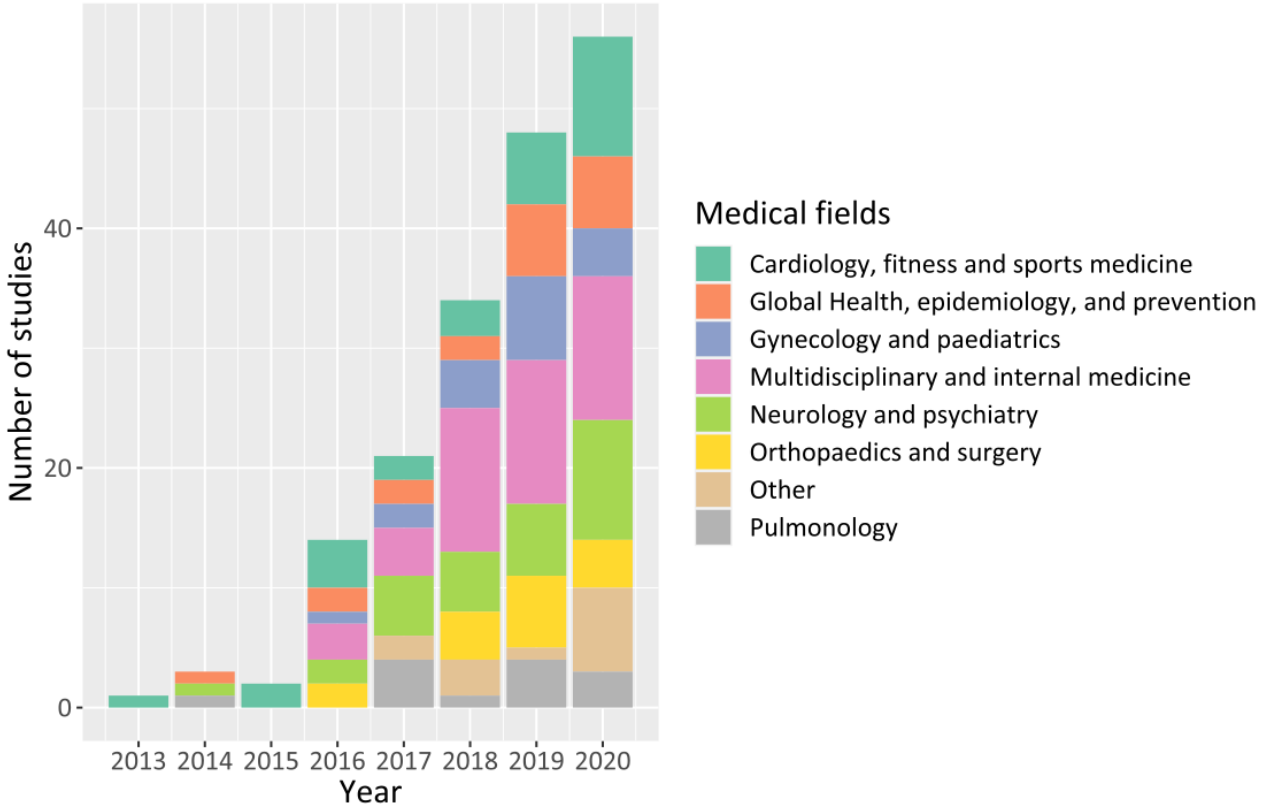
Study Types and Fields

Most studies (128/179, 71.5%) used observational study designs such as cross-sectional (66/179, 36.9%) and cohort studies (62/179, 34.6%), comprising 9,780,808 (90.26%) participants and 628,641 (5.8%) participants, out of 10,835,733 participants, respectively. Most frequently, studies (70/179, 39.1%) aimed to find associations, correlations, or influencing factors within their study population, study outcomes, and generated data. Slightly less than one-third of the studies (54/179, 30.2%) aimed to characterize and observe their study population.

Most studies were conducted in the fields of multidisciplinary and general medicine (43/179, 24%); cardiology, fitness, and sports medicine (29/179, 16.2%); and neurology, psychology, and psychiatry (28/179, 15.6%; Figure 4). The fields of global health, prevention, and epidemiology featured the largest sample size with, with 10,104,217 (93.25%) out of 10,835,733 participants.

XSL•FO
RenderX

Figure 4. Studies per medical field.



Wearable Characteristics

A total of 189 wearable devices were extracted. The company with the most wearable devices in the included studies was Fitbit (97/189, 51.3%), covering 8,361,035 (74.35%) out of 11,224,872 participants. Fitbit is followed by ActiGraph (research-grade wearable devices unavailable for consumers or not consumer grade per se; 19/189, 10.1%), Polar Electro (9/189, 4.8%), and Withings (8/189, 4.2%). In number of study participants, Huawei and Withings comprised 832,036 (7.4%) participants and 794,174 (7.06%) participants out of 11,224,872 participants, respectively (Table 2).

XSL•FO
RenderX

Table 2. Characteristics of wearable devices.

Wearable characteristics	Studies (N=189), n (%)	Participants (N=11,244,872), n (%)
Wearable companies used in studies		
Fitbit	97 (51.32)	8,361,035 (74.35)
ActiGraph[a]	19 (10.05)	2571 (0.02)
Polar electro	9 (4.76)	6970 (0.06)
Withings	8 (4.23)	794,174 (7.06)
iRhythm	6 (3.17)	128,641 (1.14)
Xiaomi	5 (2.65)	176 (<0.01)
Axivity[a]	4 (2.12)	291,871 (2.6)
Garmin	4 (2.12)	308 (<0.01)
Apple	4 (2.12)	420,826 (3.74)
Activinsights[a]	3 (1.59)	1971 (0.02)
Samsung	2 (1.06)	120 (<0.01)
Ava AG	2 (1.06)	285 (<0.01)
Huawei	2 (1.06)	832,036 (7.40)
Whoop	2 (1.06)	305 (<0.01)
Omron	2 (1.06)	159 (<0.01)
Other companies (wearable only included in 1 study)	20 (10.58)	423,424 (3.77)
Number of wearable device models per study (n=179)		
1	156 (87.15)	486,684 (4.49)
2	11 (6.15)	420,007 (3.88)
3	3 (1.68)	838,266 (7.74)
>3 or not applicable[b]	9 (5.03)	9,090,776 (83.9)
Wearable device types		
Fitness tracker	86 (45.5)	22,823 (0.2)
Accelerometer (worn on wrist, torso, and hip)	49 (25.93)	299,251 (2.66)
Electrocardiogram chest patch or strap	21 (11.11)	530,332 (4.72)
Smartwatch	12 (6.35)	1,259,605 (11.2)
Diverse wearable devices—secondary data via wearable data platform	11 (5.82)	9,122,758 (81.13)
Distinct vital sign trackers (eg, oximetry ring, temperature wristband tracker, and blood pressure armband)[c]	10 (5.29)	10,103 (0.09)
Physical location of wearable		
Wrist	138 (73.02)	10,702,843 (95.18)
Hip	25 (13.23)	2257 (0.02)
Chest	21 (11.11)	550,332 (4.89)
Arm	3 (1.59)	9392 (0.08)
Finger	2 (1.06)	48 (<0.01)
In studies used in-built sensor in wearables[d] (n=179)		
Accelerometer	146 (81.56)	1,157,069 (10.68)
Photoplethysmography	59 (32.96)	9,622,147 (88.8)
Electrodes (ie, electrocardiogram)	21 (11.73)	550,500 (5.08)
Gyroscope	6 (3.35)	1585 (0.01)

XSL•FO
RenderX

Wearable characteristics	Studies (N=189), n (%)	Participants (N=11,244,872), n (%)
Thermometer	4 (2.23)	842 (0.01)
Blood pressure sensor	3 (1.68)	9397 (0.09)
Wearable costs (€; US $)		
<200 (228)	120 (63.49)	340,460 (3.03)
200-350 (228-399)	41 (21.69)	18,256 (0.16)
>350 (399)	13 (6.88)	551,128 (4.9)
Not applicable[e]	15 (7.94)	10,355,028 (92.09)
Analysis—statistical tests[f] in studies (n=179)		
Regression	62 (34.64)	1,021,032 (9.42)
t test	41 (22.91)	8,309,202 (76.68)
Correlation (Pearson, Spearman, etc)	40 (22.35)	11,044 (0.1)
Wilcoxon *U*, Mann–Whitney *U*, and other nonparametric tests	23 (12.85)	7180 (0.07)
Chi-square and Fisher–Yates tests	15 (8.38)	433,785 (4)
Mixed methods model and other statistical models	14 (7.82)	57,938 (0.53)
Artificial Intelligence (data mining, cluster, machine learning, etc)	11 (6.15)	835,967 (7.71)
Analysis of variance	11 (6.15)	810 (0.01)
Descriptive	8 (4.47)	423,093 (3.9)
Prognostic analysis (Kaplan–Meier, permutation test, etc)	3 (1.67)	420,928 (3.88)

[a]Research-grade wearable devices unavailable for consumers or not consumer grade per se.

[b]Studies collected data with multiple wearable devices (that belonged to the study participants) or studies that used secondary data provided by web-based wearable platforms, mobile applications, or wearable companies.

[c]Distinct vital sign trackers are specialized on a specific vital sign, for example, oximetry ring, temperature wristband tracker, and blood pressure armband. They differ in measured vital signs and worn locations compared with other wearable device types.

[d]Utilized in-built sensors in wearables sums up to more than the total of wearables, as sometimes more than one built-in sensor was used.

[e]Providing wearable hardware pricing was not transparent, as some studies used data provided by diverse participant-owned wearables or wearable hardware costs were part of a subscription or a membership fee, that is, Whoop strap of Whoop.

[f]Analysis—statistical tests sums up to more than the total number of included studies, as some studies applied more than one type of analysis or statistical test.

Most studies (156/179, 87.2%) used 1 wearable model. However, most of the study participants (9,090,776/10,835,733, 83.9%) were part of large-scale population-based studies in which data were mostly collected with multiple wearable devices that belonged to the study participants.

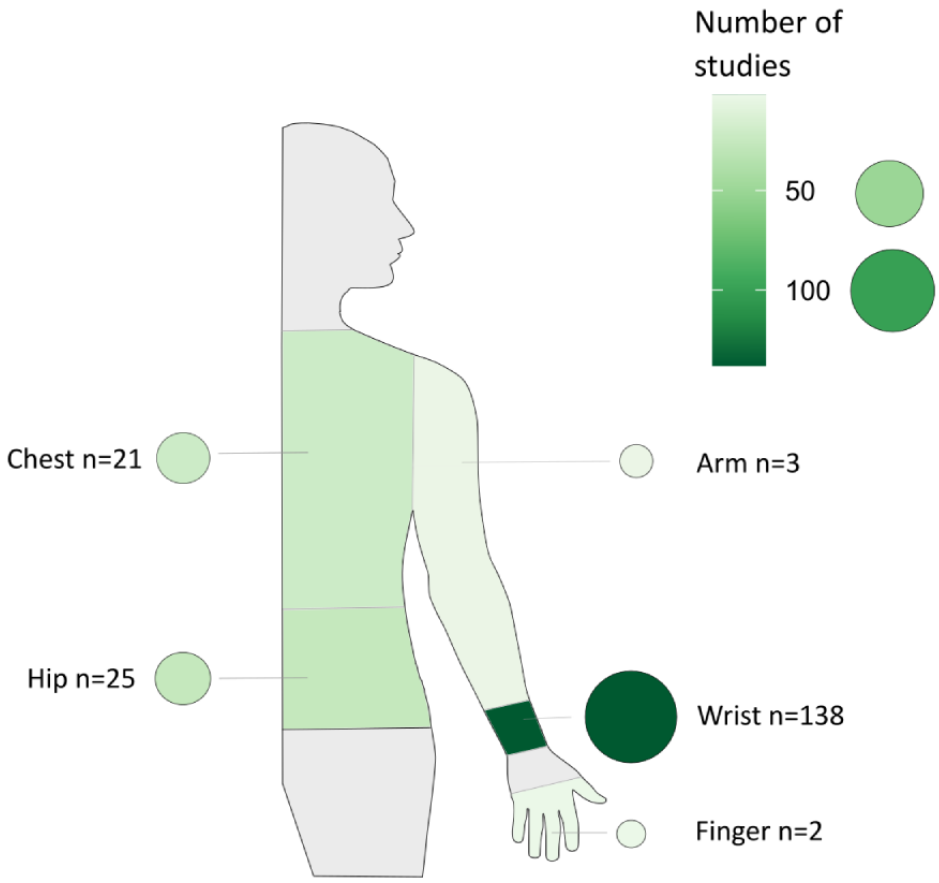
Some large-scale population-based studies (11/179, 6.1%) relied on secondary data collected with mobile apps [87] or web-based wearable platforms [153] or provided through a wearable company [189]. Thus, the device type could not be specified (assigned to category *diverse wearable devices—secondary data via wearable data platform*). A total of 15 (63%) out of 24 studies that used secondary data were conducted in 2020, and 5 (21%) studies in 2019.

Fitness trackers (86/189, 45.5%) and accelerometers (measuring body movement acceleration [37]) worn on the wrist, torso, and hip (49/189, 25.9%) were the most frequent. Other wearable device types included ECG chest straps and patches (21/189, 11.1%), smartwatches (12/189, 6.3%), and distinct vital sign trackers (10/189, 5.3%) such as oximetry rings or blood pressure armbands (Table 2).

Most wearables were worn on the wrist (138/189, 73%), followed by the hip (25/189, 13.2%) and chest (21/189, 11.1%). Only a few wearables were worn on the arm (3/189, 1.6%) and finger (2/189, 1.1%; Figure 5).

XSL•FO
RenderX

Figure 5. Wear locations of wearables and their frequencies. The color and size of the circles assigned to the body location visualize the frequency of wearables worn on the respective location.



Most of the studies used wearable built-in sensors of (1) accelerometers (146/179, 81.6%) that measure acceleration on a 3- or 1-axis [37] and (2) photoplethysmography (59/179, 33%) defined as an "optical technique that [...] detects blood volume changes in the microvascular bed of tissue" [222]. Other built-in sensors were electrodes for ECG measurements (21/179, 11.7%); gyroscopes (6/179, 3.4%), which determine how different portions of the body rotate [37]; thermometers (4/179, 2.2%) measuring skin temperature; and blood pressure sensors (3/179, 1.7%).

Most studies investigated steps (95/179, 53.1%), HR (55/179, 30.7%), and sleep time (51/179, 28.5%). We classified measured vital signs into three categories, whereby PA measures were most frequent (228/179, 127.4%; Multimedia Appendix 4 [14,15,20,44-219]):

1. PA measures included steps, intensity (eg, time spent in moderate to vigorous PA), energy expenditure (eg, kilocalories and metabolic equivalent), axial or raw movement data, distance (covered), and others (such as stairs taken, elevation, and sedentary time).
2. Cardiac measures included HR, HR variability, and ECG (or other direct heart rhythm analyses, such as AF detection).
3. Other measures that included blood or pulse pressure, body temperature, blood oxygen, and respiratory rate.

Most studies (120/189, 63.5%) used wearables that cost <€200 (US $228). In some studies (15/189, 7.9%), wearable prices were not transparent, as data were provided through a variety of participant-owned wearables [87] or the wearable hardware was part of a subscription or a membership fee, that is, Whoop strap of Whoop [178].

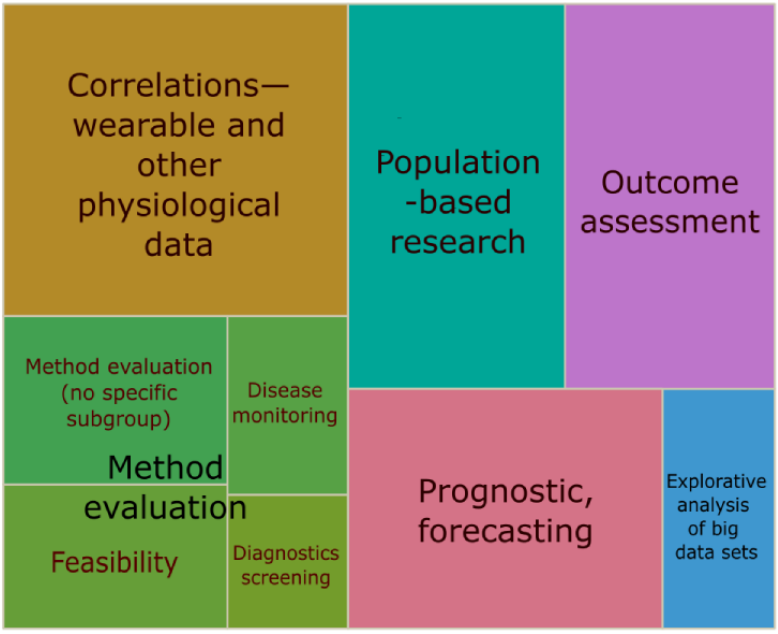
Regression analysis (62/179, 34.6%) and *t* tests (42/179, 22.9%) were the most commonly used statistical methods to analyze wearable data. Other methods comprised nonparametric tests, such as correlations, Wilcoxon U test, Kaplan-Meier survival analysis, and chi-square tests. Variance analysis (analysis of variance) and significance tests such as permutations were also used. Further data analyses were conducted in a data-driven manner [223] with artificial intelligence, such as k-means [176] or unsupervised cluster analysis [172], recursive feature elimination technique [170], rotation random forest classifier [130], and supervised machine learning algorithms using logistic regression, decision tree, and random forest [215].

Categorization of Wearable Application in the Studies

We categorized the included studies based on their study objective, the role of the wearable and the collected wearable data within the study in the following 6 categories (overlaps are possible as separation is artificial). In the following, categories are presented in order of their frequency (see Figure 6 and Multimedia Appendix 5 [14,15,20,44-219] for article references and examples).

XSL•FO
RenderX

Figure 6. Categorization of wearable applications, showing proportions of the 6 categories (with 4 subcategories). The size of depicted categories (in different colors) corresponds to the number of studies.



Correlations—Wearable and Other Physiological Data

Studies (40/179, 22.3%) have examined the correlation of a wearable derived measure with clinical- and patient-reported and other health-related outcomes to find new associations and correlations. The data generated by the wearable device were correlated with data from mostly physiological or patient-reported outcomes.

Population-Based Research

In 17.3% (31/179) of studies, wearables produced insights into a specific population through monitoring (observational and cross-sectional) of vital signs, such as steps and HR. Often, these were cross-sectional studies (17/31, 55%) where the wearable measurement was the sole outcome. The resulting data provide novel insights and characteristics of populations.

Outcome Assessment

In these studies (30/179, 17.3%), wearables generated the outcome measurement and monitored the dependent variable in an (quasi-) experimental setting or intervention, in mostly randomized controlled trials and quasi-experimental designs.

Prognosis, Forecasting, and Risk Stratification

In further studies (28/179, 15.6%), data generated with wearables were integrated into risk calculations (risk for a certain event or outcome), prognostic models, or cut-points. Wearable data constituted inputs for models to estimate risks.

Explorative Analysis of Big Data Sets

These studies (10/179, 5.6%) exploratively analyzed big data [223], generated by wearables and accessible via applications, commercial platforms, eCohorts, or companies themselves, to find trends and generate new hypotheses.

Method Evaluation

Studies (40/179, 22.3%) have evaluated and compared methods and tools (such as screenings for diseases, general practices, questionnaires, or other patient-reported outcomes) with the help of wearables. The wearable device might be the gold standard device or probed itself.

Feasibility

In these studies (12/179, 6.7%), the feasibility of using wearables for screening diseases and to improve on existing methods and practices is focused, mostly accompanied by a qualitative component.

Diagnostics and Screening

Studies (6/179, 3.4%) in this category evaluated details of diagnostics and disease screening outcomes, (cost-) effectiveness, utility, and screening length or were compared with standard measurement methods.

Disease Monitoring

Here (8/179, 4.5%), wearables supported the monitoring of an already diagnosed condition or a patient at risk (of deterioration).

Others

Studies (14/179, 7.8%) evaluated methods, with no other particular subgroup being appropriate.

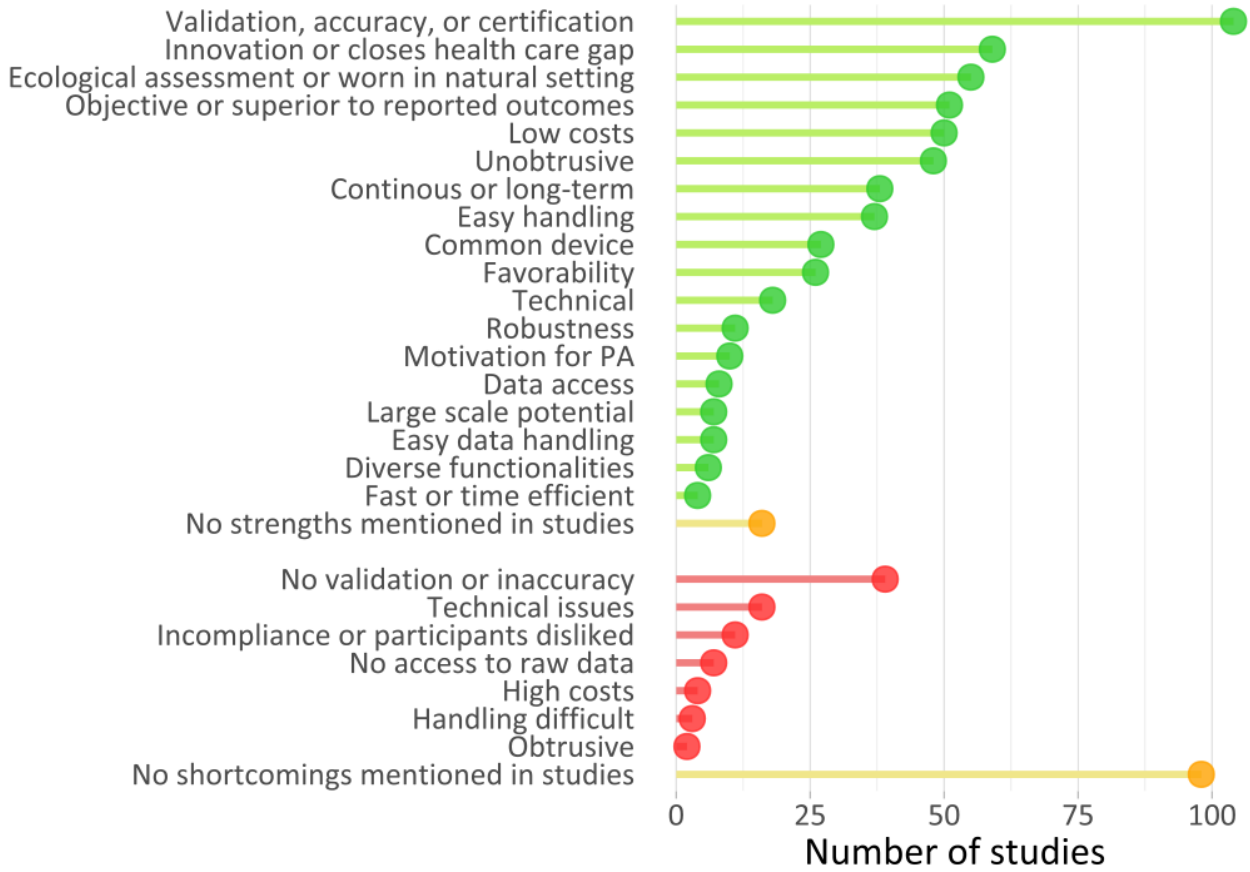
Strengths and Shortcomings of Wearables

Overall, the studies mentioned more strengths than shortcomings. A few studies (16/179, 8.9%) mentioned no strengths of wearables, whereas 55.3% (99/179) of the studies mentioned no shortcomings.

Most often, authors (104/179, 58.1%) emphasized the accuracy and reliability, positive results of peer-reviewed validation studies (own and of others), or clinically approved certifications (eg, the Food and Drug Administration [FDA] clearance in the United States or Communauté Européenne [CE] mark of the European Union; Figure 7).

XSL•FO
RenderX

Figure 7. Chart of reported strengths and weaknesses of wearables as mentioned by authors. PA: physical activity.



Often, studies (59/179, 33%) identified the wearable as innovative, that is, as a cutting-edge tool and method [103] with a wearable device potentially closing a gap in or improving health care and research. For example, 1 study described how wireless wearables and data synching could improve the quality of care [69], "The data can be sent from the wearable to the physician's office, avoiding the need for office visits, ultimately making possible preventive medicine and improving quality of care." Low et al [129] concluded that "Fitbit devices may provide opportunities to improve postoperative clinical care with minimal burden to patients or clinical providers." Tomitani et al [199] reflected how wrist-worn blood pressure wearables could "significantly improve blood pressure control." As per Shilaih et al [184], wrist-worn wearables might ameliorate fertility awareness research and care.

Several studies (55/179, 30.7%) acknowledged the ability of wearables to measure in the naturalistic environment of the participants, called *ecological momentary assessment* [10,11,224].

Multiple studies (51/179, 28.5%) described wearables as objective and superior to self-reported outcomes as they were more accurate, reliable, and easier to generate. Often, the authors valued the relatively low costs of wearables (50/179, 27.9%). Others appreciated wearables as being unobtrusive or noninvasive (48/179, 26.8%) and enabling continuous, long-term measurements (38/179, 21.2%). Furthermore, the handling (37/179, 20.7%) of hardware and software was often found to be user-friendly, as well as the prevalence of wearables in the population (27/179, 15.1%), decreasing stigma and easing participant recruitment. Some studies (26/179, 14.5%) reported that participants accepted and liked the wearables, resulting in high participant compliance (wearing and using the wearable). Some authors (18/179, 10.1%) perceived technical wearable characteristics as positive, for example, good sampling rate of measurements, long battery life, large memory space, raw data availability, data security, compatibility with other devices such as smartphones, and availability of application programing interfaces (APIs).

Few studies (11/179, 6.1%) described wearables as robust and not easy to break. Authors (10/179, 5.6%) valued the wearable-induced behavior change as a cobenefit, that is, motivating study participants to more PA and increasing health awareness.

A few studies (8/179, 4.5%) mentioned data accessibility via APIs, apps, and web-based platforms and a few other studies (7/179, 3.9%) potential of large-scale wearable studies, or the ease of data handling. A few (6/179, 3.4%) studies underlined the variety of functionalities and vital sign measurements as positive aspects, and 2.2% (4/179) of studies perceived wearables as fast or time-efficient in data generation.

Most shortcomings (39/179, 21.8%) were related to the inaccuracy of the wearables or the absence of validation or clinically approved certification. Studies (16/179, 8.9%) also mentioned technical issues, such as a low sampling rate of measurements, no wear time recognition, or missing data. Other technical issues comprised, for example, synchronization,

XSL•FO
RenderX

charging and device setup [91] or data cleaning [137]. Rare experienced shortcomings were participants' noncompliance or dislike toward the wearable (11/179, 6.1%), no access to raw data or company's algorithms (4/179, 2.2%), difficulties in handling the wearable (3/179, 1.7%), and wearables perceived as obtrusive in daily life (2/179, 1.1%).

Discussion

Study Characteristics

Overall, we have identified a positive trend in wearable studies, underlining the growing interest in wearables in health research, in line with other reviews [3,224-226]. Our results show a strong interest of researchers and study participants in this technology, but we also identified cautionary behavior toward using wearables. The vast majority of studies were undertaken in North America, about twice as many as in Europe, which is consistent with the previous literature [225]. One study in North America, conducted in 2019 with over 8 million participants [153], dominated the image of the distribution of participants. The reasons for the American-European gap may be multifaceted. One factor may be the differences in political and administrative frameworks, for example, comparing CE and FDA processes, which may result in slower certification processes for wearables and new technologies in general [31]. Another factor may be cultural mentality resulting in faster adoption of new technology in the United States, as the North Americans own proportionally more fitness trackers in comparison to the Europeans [227,228].

Some factors discussed in other research were not or only briefly mentioned in the included studies [6,29,31], but should also be reflected, especially technical and legal aspects, such as data security [224], data synching, and export. For example, the Germany-based study of Koehler et al [114] was one of the few that detailed data security and transfer of home-based telemonitoring data to the clinic. Data security and privacy are severely governed by the European Union General Data Protection Regulation, which is according to their website the "toughest privacy and security law in the world" [229]. Administrative limitations and challenges presumably obscure the benefits of wearable research in Europe. A possible solution for data security and usability might be data trusts [230] as an alternative to large platforms.

Most medical fields represented in the included studies showed similarities with other reviews [224], for example, studies often focusing on cardiology, sports medicine, and neurology. However, we found a multitude of studies from multidisciplinary fields as well as the field of global health, indicating a likely adoption and expansion of wearables in other medical fields. This underlines the potential for wearables in health research beyond a mere trend or hype, as wearables may provide new possibilities for a broad spectrum of health research, such as for infectious disease prediction like COVID-19 or fertility awareness, among many others.

Wearable Characteristics

Similar to other reviews, most devices were wrist-worn fitness trackers and accelerometers, and most of them are from the company Fitbit, measuring PA, HR, and sleep [3,27,31,224,225]. These vital signs and device types seem to become the standard in wearable research [3,27,31,224,225]. The included studies also emphasized the growing wearable use [147,195,197], which is also reflected in commercially available devices [38-40]. Currently, further wearable devices emerge, measuring, for example, oximetry, blood pressure, skin temperature, or respiratory rate.

Categorization of Wearable Application in the Studies

In general, the included studies covered a great scope of health applications such as fertility tracking; monitoring of body characteristics such as weight or diseases such as Alzheimer disease, diabetes mellitus, and AF; as well as associations of coffee intake, sleep, and PA, or blood pressure and steps. We have noted an increase in smaller studies that also included rare populations and conditions, such as fibromyalgia or the rare genetic Pompe disease, indicating that wearables may be valuable for insights into patients with rare conditions. Using affordable, consumer-grade wearables for rare disease assessment and monitoring might eventually be less expensive than specifically developed devices and easier to use for patients. Therefore, currently underrepresented populations may be better researched through wearables [231], that is, different ethnic groups, nationalities, individuals with disabilities, or (rare) conditions. Future studies could examine the participation of underrepresented groups in wearable research in greater depth, particularly in studies analyzing wearable user data.

Global Health and Low-Resource Contexts

Included studies are predominantly from high-income countries, constituting a gap in wearable studies in low-resource contexts. The AliveCor device was shown to be feasible in Kenya to help detect AF [232], as well as for early diagnosis. The literature underlines the potential for wearable-based research in low-resource settings to generate data and improve health care [9], based on their low cost and ease of use (data acquisition, hardware, and software handling) [233]. Xu et al [234] emphasized that physiological monitoring with wearables hold "promise for substantial improvements in neonatal outcomes" in low- and middle-resource countries. Wearables can generate a solid database for global health research, particularly for morbidity measurements [235], large-scale studies, and modeling and descriptive studies. Topics such as climate change–induced impacts focusing on extreme weather events as an outcome and impact on health [236] may be approached. For example, 1 study [20] measured the physiological response of farm workers to climate conditions with wearables to investigate heat-related illness in a high-income setting. Lam et al [237] investigated the thermal adaptation and comfort of participants originating from various climatic regions. The fitness tracker measured HR data was integrated with other weather and human-based measurements and predicted the thermal sensation of nonlocal participants, among others. Similar studies can be conducted in low-resource regions.

Strengths and Shortcomings of Wearables

A few studies have experienced issues or shortcomings, such as inaccuracies in measurements and technical issues.

XSL•FO
RenderX

Nevertheless, most authors were satisfied with wearables, as strengths were mentioned more frequently than shortcomings. Novelty and innovation outweighed the shortcomings for most authors. The most mentioned positive wearable characteristics were validity and accuracy, technical reliability, innovation, and unobtrusiveness. Only a few authors have mentioned data access through APIs or cloud platforms as a strength. However, the practical value of wearables is heavily reliant on the mode and reliability of data access. Depending on the company, there may be different data access policies in place, whereby it may not be possible to access the raw data of the wearable. Most authors have not considered wearable data access. However, data access and availability of wearable devices is an important aspect that researchers need to be aware of before using a potential study device. Another aspect is open access to the wearables' raw data or source codes, as companies might change the source code and implement algorithms without the obligation to announce or detail changes that might lead to bias and inconsistency of data [224]. For example, Thijs et al [195] mentioned the consequences of nondisclosed algorithms (Fitbit) for data analysis and standardization. Moreover, the lack of standardization and replicability of wearable raw data and analysis [28] hinders comparability among studies.

Most studies mentioned and discussed validation, accuracy, and certification of the used wearables as part of good research practice approaches. However, the mention of validation or accuracy did not necessarily imply that the wearables had been certified (FDA or CE) or validated in peer-reviewed research. Nevertheless, the authors reported that the wearable device is sufficiently accurate even with existing inaccuracies [14,143,197]. The authors seemed to tolerate smaller inaccuracies and validation drawbacks—especially of established consumer-grade wearables—if usability was of high importance, such as in large-scale studies.

Large-scale and Big Data Sets for Wearable Research

We noted an increase in large-scale wearable studies in recent years, which is consistent with previous literature [225]. During the COVID-19 pandemic, there has been an increase in studies using secondary data. Studies aimed at generating insights with regard to the developing pandemic, focusing on forecasting models and their effects on different populations. Overall, wearable-generated big data sets might decrease biased data because measurements are objectively taken in the natural environment of numerous and diverse individuals. Although data analytic skills are needed for handling big data sets, their analysis might be extremely valuable for health research in generating new evidence [31,225].

Limitations

First, not all studies using wearables might have been identified by our search. We included only the wearables of companies in the search that had the highest market share. Therefore, the wearables of smaller or new companies may be missing in this review. In addition, we only included studies published in English, which may have excluded evidence from other regions that may not publish in English. Although this review provides a wide scope of wearable research, the list of included studies is by no means exhaustive.

In addition, wearable costs are only approximations and could be imprecise: (1) companies follow different sales and distribution models, for example membership, rental, and subscription; (2) we only incorporated wearable (hardware) prices, excluding costs for software, maintenance, and other charges such as subscription fees, which may even exceed wearable hardware costs; *and* (3) sales prices are subject to fluctuation. We also excluded many studies as wearables were discontinued. The fluctuant and unstable market, therefore, might also be a factor in decisions regarding the use of wearables [28]. Although interesting and promising, some wearables and similar devices were beyond the scope of this study but might also be valuable for health research. We have provided a wide overview of wearable devices; however, the included studies did not show the full range of possible wearables and measured vital signs [9,37].

In addition, we report the opinions of the included studies with regard to the shortcomings and strengths of wearables. Although these insights might be helpful, they are not objective measures. Moreover, our introduced categories for studies and aims to use wearables might overlap, as separation and categorization are artificial.

Conclusions

We see a growing uptake of wearables in health research and a trend to use wearables for large-scale, population-based studies. Wearables, which were often piloted in the included studies, were used in diverse health fields including COVID-19 prediction, fertility awareness, geriatrics, AF detection, evaluation of methods, drug effects, psychological interventions, and patient-reported outcomes. Measurement of steps, PA, HR, and sleep may be considered standard wearable measurements. Nevertheless, wearables are becoming more diverse in their measurements and appearance. Therefore, wearable-induced research may include currently underrepresented populations such as the older adults, participants who are disabled, participants with rare chronic or genetic diseases, participants from low socioeconomic backgrounds, and others.

For many researchers, novelty and innovation seem to outweigh shortcomings such as measurement inaccuracies. Overall, the included studies shared key characteristics that the wearables should meet: validity, technical reliability (including data access solutions), innovation, and unobtrusiveness.

We identified a lack of wearable research in low-resource settings. We assume that the reasons for the gap may be a lack of funding and doubts about the usefulness of the wearables. However, wearable devices may be used to generate data in such settings, which may otherwise be difficult and expensive to obtain. Therefore, wearable devices may be valuable for health research in a global context. During the COVID-19-pandemic in particular, large-sized wearable studies were used to generate insights into the developing pandemic and may potentially lead to novel insights into population health trends and forecasts. Future research is needed to determine the usability of wearable devices for underrepresented populations, as well as the feasibility and usefulness of health research in low-resource contexts.

XSL•FO
RenderX

Acknowledgments

We wish to thank the German Research Foundation (Deutsche Forschungsgemeinschaft) for supporting this study as part of a Deutsche Forschungsgemeinschaft–funded research unit (Forschungsgruppe). We acknowledge the support of Else Kröner-Fresenius-Stiftung from the Heidelberg Graduate School of Global Health. Funders did not have a role in the design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

SH, SB, and MA conceived and designed the study. SH drafted the manuscript with the help of SB and MA. All authors contributed to the critical revision of the draft and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[DOCX File , 116 KB - mhealth_v10i1e34384_app1.docx]

Multimedia Appendix 2
Details on search and search strings.
[DOCX File , 16 KB - mhealth_v10i1e34384_app2.docx]

Multimedia Appendix 3
Medical Education Research Study Quality Instrument scores of included studies.
[DOCX File , 62 KB - mhealth_v10i1e34384_app3.docx]

Multimedia Appendix 4
Vital signs measured by studies.
[DOCX File , 254 KB - mhealth_v10i1e34384_app4.docx]

Multimedia Appendix 5
Categorization of wearable applications in the studies: article references and examples.
[DOCX File , 65 KB - mhealth_v10i1e34384_app5.docx]

References

1. Wearable band shipments rocket by 684%. Canalys. Singapore; 2014. URL: https://www.canalys.com/newsroom/wearable-band-shipments-rocket-684?time=1605182466 [accessed 2020-11-12]
2. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
3. Shin G, Jarrahi MH, Fei Y, Karami A, Gafinowitz N, Byun A, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: a systematic literature review. J Biomed Inform 2019;93:103153. [doi: 10.1016/j.jbi.2019.103153] [Medline: 30910623]
4. Carpenter A, Frontera A. Smart-watches: a potential challenger to the implantable loop recorder. Europace 2016;18(6):791-793. [doi: 10.1093/europace/euv427] [Medline: 26847074]
5. Jia Y, Wang W, Wen D, Liang L, Gao L, Lei J. Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. PeerJ 2018;6:e5350 [FREE Full text] [doi: 10.7717/peerj.5350] [Medline: 30065893]
6. Hickey AM, Freedson PS. Utility of consumer physical activity trackers as an intervention tool in cardiovascular disease prevention and treatment. Prog Cardiovasc Dis 2016;58(6):613-619. [doi: 10.1016/j.pcad.2016.02.006] [Medline: 26943981]
7. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
8. Zhang Y, Weaver RG, Armstrong B, Burkart S, Zhang S, Beets MW. Validity of wrist-worn photoplethysmography devices to measure heart rate: a systematic review and meta-analysis. J Sports Sci 2020;38(17):2021-2034. [doi: 10.1080/02640414.2020.1767348] [Medline: 32552580]

XSL•FO
RenderX

9. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Per Med 2018;15(5):429-448 [FREE Full text] [doi: 10.2217/pme-2018-0044] [Medline: 30259801]
10. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
11. Zapata-Lamana R, Lalanza JF, Losilla JM, Parrado E, Capdevila L. mHealth technology for ecological momentary assessment in physical activity research: a systematic review. PeerJ 2020;8:e8848 [FREE Full text] [doi: 10.7717/peerj.8848] [Medline: 32257648]
12. Berrut G. La montre connectée : un instrument de santé publique? Geriatr Psychol Neuropsychiatr Vieil 2020;18(1):5. [doi: 10.1684/pnv.2020.0847] [Medline: 32160985]
13. Robert Koch Institute. Corona-Datenspende | Robert Koch-Institut. Corona Data Donation. 2020. URL: https://corona-datenspende.de/ [accessed 2020-10-26]
14. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digit Health 2020;2(2):e85-e93 [FREE Full text] [doi: 10.1016/S2589-7500(19)30222-5] [Medline: 33334565]
15. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019;381(20):1909-1917. [doi: 10.1056/NEJMoa1901183] [Medline: 31722151]
16. Worsham C, Jena AB. Why doctors shouldn't dismiss the Apple watch's new ECG app. Harvard Business Review. 2018. URL: https://hbr.org/2018/10/why-doctors-shouldnt-dismiss-the-apple-watchs-new-ecg-app [accessed 2020-10-02]
17. Dörr M, Nohturfft V, Brasier N, Bosshard E, Djurdjevic A, Gross S, et al. Reply: watch out: the many limitations in smartwatch-driven AF detection. JACC Clin Electrophysiol 2019;5(4):526-527 [FREE Full text] [doi: 10.1016/j.jacep.2019.03.001] [Medline: 31000111]
18. Kaneshiro M, Kaiser W, Pourmorady J, Fleshner P, Russell M, Zaghiyan K, et al. Postoperative gastrointestinal telemetry with an acoustic biosensor predicts ileus vs. uneventful GI recovery. J Gastrointest Surg 2016;20(1):132-139. [doi: 10.1007/s11605-015-2956-3] [Medline: 26408329]
19. Barajas-Carmona JG, Francisco-Aldana L, Morales-Narváez E. Wearable nanoplasmonic patch detecting Sun/UV exposure. Anal Chem 2017;89(24):13589-13595. [doi: 10.1021/acs.analchem.7b04066] [Medline: 29155557]
20. Culp K, Tonelli S. Heat-related illness in midwestern hispanic farmworkers: a descriptive analysis of hydration status and reported symptoms. Workplace Health Saf 2019;67(4):168-178 [FREE Full text] [doi: 10.1177/2165079918813380] [Medline: 30724664]
21. Choi DH, Thaxton A, Jeong IC, Kim K, Sosnay PR, Cutting GR, et al. Sweat test for cystic fibrosis: wearable sweat sensor vs. standard laboratory test. J Cyst Fibros 2018;17(4):e35-e38 [FREE Full text] [doi: 10.1016/j.jcf.2018.03.005] [Medline: 29580829]
22. Relf R, Willmott A, Flint MS, Beale L, Maxwell N. Reliability of a wearable sweat rate monitor and routine sweat analysis techniques under heat stress in females. J Therm Biol 2019;79:209-217. [doi: 10.1016/j.jtherbio.2018.12.019] [Medline: 30612681]
23. Kwon S, Hong J, Choi EK, Lee B, Baik C, Lee E, et al. Detection of atrial fibrillation using a ring-type wearable device (CardioTracker) and deep learning analysis of photoplethysmography signals: prospective observational proof-of-concept study. J Med Internet Res 2020;22(5):e16443 [FREE Full text] [doi: 10.2196/16443] [Medline: 32348254]
24. Genbrugge C, Eertmans W, Salcido DD. Monitor the quality of cardiopulmonary resuscitation in 2020. Curr Opin Crit Care 2020;26(3):219-227. [doi: 10.1097/MCC.0000000000000726] [Medline: 32332284]
25. Caduff A, Zanon M, Zakharov P, Mueller M, Talary M, Krebs A, et al. First experiences with a wearable multisensor in an outpatient glucose monitoring study, part I: the users' view. J Diabetes Sci Technol 2018;12(3):562-568 [FREE Full text] [doi: 10.1177/1932296817750932] [Medline: 29332423]
26. van der Kamp MR, Klaver EC, Thio BJ, Driessen JM, de Jongh FH, Tabak M, et al. WEARCON: wearable home monitoring in children with asthma reveals a strong association with hospital based assessment of asthma control. BMC Med Inform Decis Mak 2020;20(1):192 [FREE Full text] [doi: 10.1186/s12911-020-01210-1] [Medline: 32795352]
27. Foster KR, Torous J. The opportunity and obstacles for smartwatches and wearable sensors. IEEE Pulse 2019;10(1):22-25. [doi: 10.1109/MPULS.2018.2885832] [Medline: 30872210]
28. Shelgikar AV, Anderson PF, Stephens MR. Sleep tracking, wearable technology, and opportunities for research and clinical care. Chest 2016;150(3):732-743. [doi: 10.1016/j.chest.2016.04.016] [Medline: 27132701]
29. Jin D, Adams H, Cocco AM, Martin WG, Palmer S. Smartphones and wearable technology: benefits and concerns in cardiology. Med J Aust 2020;212(2):54-6.e1. [doi: 10.5694/mja2.50446] [Medline: 31851375]
30. Haberlin C, O'Dwyer T, Mockler D, Moran J, O'Donnell DM, Broderick J. The use of eHealth to promote physical activity in cancer survivors: a systematic review. Support Care Cancer 2018;26(10):3323-3336. [doi: 10.1007/s00520-018-4305-z] [Medline: 29909476]
31. Jacobsen M, Dembek TA, Kobbe G, Gaidzik PW, Heinemann L. Noninvasive continuous monitoring of vital signs with wearables: fit for medical use? J Diabetes Sci Technol 2021;15(1):34-43 [FREE Full text] [doi: 10.1177/1932296820904947] [Medline: 32063034]

XSL•FO
RenderX

32. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
33. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
34. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
35. Elliott M, Coventry A. Critical care: the eight vital signs of patient monitoring. Br J Nurs 2012;21(10):621-625. [doi: 10.12968/bjon.2012.21.10.621] [Medline: 22875303]
36. Dias D, Paulo Silva Cunha J. Wearable health devices-vital sign monitoring, systems and technologies. Sensors (Basel) 2018;18(8):2414 [FREE Full text] [doi: 10.3390/s18082414] [Medline: 30044415]
37. Kamišalić A, Fister Jr I, Turkanović M, Karakatič S. Sensors and functionalities of non-invasive wrist-wearable devices: a review. Sensors (Basel) 2018;18(6):1714 [FREE Full text] [doi: 10.3390/s18061714] [Medline: 29799504]
38. Fortune Business Insights. Fitness tracker market to reach USD 91.98 billion by 2027; rising health disorders to brighten market prospects, states fortune business insights™. GlobeNewswire. 2020. URL: https://www.globenewswire.com/news-release/2020/08/03/2071561/0/en/Fitness-Tracker-Market-to-Reach-USD-91-98-Billion-by-2027-Rising-Health-Disorders- to-Brighten-Market-Prospects-states-Fortune-Business-Insights.html [accessed 2021-07-13]
39. International Data Corporation. Wearable devices market share. IDC Corporate USA. 2021. URL: https://www.idc.com/promo/wearablevendor/vendor [accessed 2021-07-13]
40. Vailshery LS. Wearables shipments worldwide market share 2014-2021, by vendor. Statista. 2021. URL: https://www.statista.com/statistics/435944/quarterly-wearables-shipments-worldwide-market-share-by-vendor/ [accessed 2021-07-13]
41. Covidence systematic review software. Covidence. Melbourne: Veritas Health Innovation URL: https://www.covidence.org/ [accessed 2021-03-15]
42. Zotero. Roy Rosenzweig Center for History and New Media. Fairfax: Roy Rosenzweig Center for History and New Media URL: https://www.zotero.org/download/ [accessed 2021-03-15]
43. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the newcastle-ottawa scale-education. Acad Med 2015;90(8):1067-1076. [doi: 10.1097/ACM.0000000000000786] [Medline: 26107881]
44. Agarwal DK, Viers BR, Rivera ME, Nienow DA, Frank I, Tollefson MK, et al. Physical activity monitors can be successfully implemented to assess perioperative activity in urologic surgery. Mhealth 2018;4:43 [FREE Full text] [doi: 10.21037/mhealth.2018.09.05] [Medline: 30363722]
45. Alghamdi AS, Alghamdi KA, Jenkins RO, Alghamdi MN, Haris PI. Impact of Ramadan on physical activity and sleeping patterns in individuals with type 2 diabetes: the first study using fitbit device. Diabetes Ther 2020;11(6):1331-1346 [FREE Full text] [doi: 10.1007/s13300-020-00825-x] [Medline: 32367477]
46. An KH, Han KA, Sohn TS, Park IB, Kim HJ, Moon SD, et al. Body fat is related to sedentary behavior and light physical activity but not to moderate-vigorous physical activity in type 2 diabetes mellitus. Diabetes Metab J 2020;44(2):316-325 [FREE Full text] [doi: 10.4093/dmj.2019.0029] [Medline: 31769237]
47. Arigo D, Pasko K, Mogle JA. Daily relations between social perceptions and physical activity among college women. Psychol Sport Exerc 2020;47:101528 [FREE Full text] [doi: 10.1016/j.psychsport.2019.04.018] [Medline: 32831642]
48. Bade BC, Brooks MC, Nietert SB, Ulmer A, Thomas DD, Nietert PJ, et al. Assessing the correlation between physical activity and quality of life in advanced lung cancer. Integr Cancer Ther 2018;17(1):73-79 [FREE Full text] [doi: 10.1177/1534735416684016] [Medline: 28024420]
49. Baril JF, Bromberg S, Moayedi Y, Taati B, Manlhiot C, Ross HJ, et al. Use of free-living step count monitoring for heart failure functional classification: validation study. JMIR Cardio 2019;3(1):e12122 [FREE Full text] [doi: 10.2196/12122] [Medline: 31758777]
50. Barkley R, Khalil M, Shen P, Levine EA, Votanopoulos K, Clark CJ. Feasibility of low-cost accelerometers in measuring functional recovery after major oncologic surgery. J Surg Oncol 2019;121(2):279-285. [doi: 10.1002/jso.25789] [Medline: 31782174]
51. Barrett PM, Komatireddy R, Haaser S, Topol S, Sheard J, Encinas J, et al. Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. Am J Med 2014;127(1):95.e11-95.e17 [FREE Full text] [doi: 10.1016/j.amjmed.2013.10.003] [Medline: 24384108]
52. Bevier W, Glantz N, Hoppe C, Morales Glass J, Larez A, Chen K, et al. Self-reported and objectively measured physical activity levels among Hispanic/Latino adults with type 2 diabetes. BMJ Open Diabetes Res Care 2020;8(1):e000893 [FREE Full text] [doi: 10.1136/bmjdrc-2019-000893] [Medline: 32169933]
53. Bian J, Guo Y, Xie M, Parish AE, Wardlaw I, Brown R, et al. Exploring the association between self-reported asthma impact and fitbit-derived sleep quality and physical activity measures in adolescents. JMIR Mhealth Uhealth 2017;5(7):e105 [FREE Full text] [doi: 10.2196/mhealth.7346] [Medline: 28743679]

XSL•FO
RenderX

54. Billeci L, Tonacci A, Narzisi A, Manigrasso Z, Varanini M, Fulceri F, et al. Heart rate variability during a joint attention task in toddlers with autism spectrum disorders. Front Physiol 2018;9:467 [FREE Full text] [doi: 10.3389/fphys.2018.00467] [Medline: 29765335]
55. Birkeland K, Khandwalla RM, Kedan I, Shufelt CL, Mehta PK, Minissian MB, et al. Daily activity measured with wearable technology as a novel measurement of treatment effect in patients with coronary microvascular dysfunction: substudy of a randomized controlled crossover trial. JMIR Res Protoc 2017;6(12):e255 [FREE Full text] [doi: 10.2196/resprot.8057] [Medline: 29263019]
56. Blackshear TB, Seyfried L. Does education close the black-white physical activity and obesity gaps? J Am Coll Health 2021;69(2):222-226. [doi: 10.1080/07448481.2019.1657122] [Medline: 31765289]
57. Block VJ, Bove R, Zhao C, Garcha P, Graves J, Romeo AR, et al. Association of continuous assessment of step count by remote monitoring with disability progression among adults with multiple sclerosis. JAMA Netw Open 2019;2(3):e190570 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.0570] [Medline: 30874777]
58. Block VJ, Lizée A, Crabtree-Hartman E, Bevan CJ, Graves JS, Bove R, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. J Neurol 2017;264(2):316-326. [doi: 10.1007/s00415-016-8334-6] [Medline: 27896433]
59. Bolourchi M, Batra AS. Diagnostic yield of patch ambulatory electrocardiogram monitoring in children (from a national registry). Am J Cardiol 2015;115(5):630-634. [doi: 10.1016/j.amjcard.2014.12.014] [Medline: 25591894]
60. Alvarez MP, Silva TD, Favero FM, Valenti VE, Raimundo RD, Vanderlei LC, et al. Autonomic modulation in Duchenne muscular dystrophy during a computer task: a prospective control trial. PLoS One 2017;12(1):e0169633 [FREE Full text] [doi: 10.1371/journal.pone.0169633] [Medline: 28118369]
61. Brazendale K, Brazendale AB, Garcia JM, Monroe CM, Weaver RG, Beets MW. Brief report: obesogenic behaviors of children with developmental disabilities during summer. J Autism Dev Disord 2021;51(2):734-740. [doi: 10.1007/s10803-020-04566-5] [Medline: 32533383]
62. Brudy L, Hock J, Häcker AL, Meyer M, Oberhoffer R, Hager A, et al. Children with congenital heart disease are active but need to keep moving: a cross-sectional study using wrist-worn physical activity trackers. J Pediatr 2020;217:13-19. [doi: 10.1016/j.jpeds.2019.09.077] [Medline: 31740142]
63. Buchan DS, McLellan G, Donnelly S, Arthur R. The use of the intensity gradient and average acceleration metrics to explore associations with BMI z-score in children. J Sports Sci 2019;37(23):2751-2758. [doi: 10.1080/02640414.2019.1664536] [Medline: 31506039]
64. Burnett-Zeigler IE, Waldron EM, Hong S, Yang A, Wisner KL, Ciolino JD. Accessibility and feasibility of using technology to support mindfulness practice, reduce stress and promote long term mental health. Complement Ther Clin Pract 2018;33:93-99. [doi: 10.1016/j.ctcp.2018.09.001] [Medline: 30396633]
65. Cai M, Tan KH, Ang SB. I-ACT: integrated study on effect of activity on complications in pregnancy: study protocol of a multiethnic prospective cohort study. BMJ Open 2019;9(4):e025970 [FREE Full text] [doi: 10.1136/bmjopen-2018-025970] [Medline: 30948592]
66. Carrasco JJ, Pérez-Alenda S, Casaña J, Soria-Olivas E, Bonanad S, Querol F. Physical activity monitoring and acceptance of a commercial activity tracker in adult patients with haemophilia. Int J Environ Res Public Health 2019;16(20):3851 [FREE Full text] [doi: 10.3390/ijerph16203851] [Medline: 31614706]
67. Cascino TM, McLaughlin VV, Richardson CR, Behbahani-Nejad N, Moles VM, Visovatti SH, et al. Barriers to physical activity in patients with pulmonary hypertension. Pulm Circ 2019;9(2):2045894019847895 [FREE Full text] [doi: 10.1177/2045894019847895] [Medline: 30983524]
68. Chang YT. Physical activity and cognitive function in mild cognitive impairment. ASN Neuro 2020;12:1759091419901182 [FREE Full text] [doi: 10.1177/1759091419901182] [Medline: 31948261]
69. Chapple C, Bliwise D, Maislisch L, Roitmann E, Burtea T. Night-time voids, level of bother and sleep characteristics in a non-patient population of wearable devices users. Int J Clin Pract 2020;74(7):e13495 [FREE Full text] [doi: 10.1111/ijcp.13495] [Medline: 32100396]
70. Chue AE, Gunthert KC, Kim RW, Alfano CA, Ruggiero AR. The role of sleep in adolescents' daily stress recovery: negative affect spillover and positive affect bounce-back effects. J Adolesc 2018;66:101-111. [doi: 10.1016/j.adolescence.2018.05.006] [Medline: 29842996]
71. Cohen-Holzer M, Sorek G, Schweizer M, Katz-Leurer M. The influence of a constraint and bimanual training program using a variety of modalities on endurance and on the cardiac autonomic regulation system of children with unilateral cerebral palsy: a self-control clinical trial. NeuroRehabilitation 2017;41(1):119-126. [doi: 10.3233/NRE-171463] [Medline: 28505992]
72. Cole TS, Jahnke H, Godzik J, Morgan CD, Nakaji P, Little AS. Use of a wrist-mounted device for continuous outpatient physiologic monitoring after transsphenoidal surgery: a pilot study. Pituitary 2019;22(2):156-162. [doi: 10.1007/s11102-019-00946-y] [Medline: 30806859]
73. Collier E, Varon C, Van Huffel S, Bogaert G. Enuretic children have a higher variability in REM sleep when comparing their sleep parameters with nonenuretic control children using a wearable sleep tracker at home. Neurourol Urodyn 2020;39(1):367-375. [doi: 10.1002/nau.24215] [Medline: 31729062]

XSL•FO
RenderX

74. Costa J, Figueiredo P, Nakamura F, Rago V, Rebelo A, Brito J. Intra-individual variability of sleep and nocturnal cardiac autonomic activity in elite female soccer players during an international tournament. PLoS One 2019;14(9):e0218635 [FREE Full text] [doi: 10.1371/journal.pone.0218635] [Medline: 31527865]
75. Currow D, Watts GJ, Johnson M, McDonald CF, Miners JO, Somogyi AA, Australian national Palliative Care Clinical Studies Collaborative (PaCCSC). A pragmatic, phase III, multisite, double-blind, placebo-controlled, parallel-arm, dose increment randomised trial of regular, low-dose extended-release morphine for chronic breathlessness: Breathlessness, Exertion And Morphine Sulfate (BEAMS) study protocol. BMJ Open 2017;7(7):e018100 [FREE Full text] [doi: 10.1136/bmjopen-2017-018100] [Medline: 28716797]
76. DasMahapatra P, Chiauzzi E, Bhalerao R, Rhodes J. Free-living physical activity monitoring in adult US patients with multiple sclerosis using a consumer wearable device. Digit Biomark 2018;2(1):47-63 [FREE Full text] [doi: 10.1159/000488040] [Medline: 32095756]
77. Do J, Webster RJ, Longmuir PE, Ieradi S, Reddy D, Whiting S, et al. Physically active children with epilepsy have good objective sleep duration and efficiency despite subjective reports of fatigue and sleep problems. Epilepsy Behav 2020;104(Pt A):106853. [doi: 10.1016/j.yebeh.2019.106853] [Medline: 31958642]
78. Downey C, Randell R, Brown J, Jayne DG. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018;20(12):e10802 [FREE Full text] [doi: 10.2196/10802] [Medline: 30538086]
79. Driesman AS, Strauss EJ, Konda SR, Egol KA. Factors associated with orthopaedic resident burnout: a pilot study. J Am Acad Orthop Surg 2020;28(21):900-906. [doi: 10.5435/JAAOS-D-19-00648] [Medline: 32039922]
80. Elmagboul N, Coburn BW, Foster J, Mudano A, Melnick J, Bergman D, et al. Physical activity measured using wearable activity tracking devices associated with gout flares. Arthritis Res Ther 2020;22(1):181 [FREE Full text] [doi: 10.1186/s13075-020-02272-2] [Medline: 32746893]
81. English C, Healy GN, Coates A, Lewis L, Olds T, Bernhardt J. Sitting and activity time in people with stroke. Phys Ther 2016;96(2):193-201. [doi: 10.2522/ptj.20140522] [Medline: 26112254]
82. Eyre EL, Duncan MJ, Birch SL, Cox V, Blackett M. Physical activity patterns of ethnic children from low socio-economic environments within the UK. J Sports Sci 2015;33(3):232-242. [doi: 10.1080/02640414.2014.934706] [Medline: 24998418]
83. Fagherazzi G, El Fatouhi D, Bellicha A, El Gareh A, Affret A, Dow C, et al. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. J Med Internet Res 2017;19(10):e363 [FREE Full text] [doi: 10.2196/jmir.7930] [Medline: 29061551]
84. Fairclough SJ, Taylor S, Rowlands AV, Boddy LM, Noonan RJ. Average acceleration and intensity gradient of primary school children and associations with indicators of health and well-being. J Sports Sci 2019;37(18):2159-2167. [doi: 10.1080/02640414.2019.1624313] [Medline: 31156048]
85. Faust L, Feldman K, Mattingly SM, Hachen D, V Chawla N. Deviations from normal bedtimes are associated with short-term increases in resting heart rate. NPJ Digit Med 2020;3:39 [FREE Full text] [doi: 10.1038/s41746-020-0250-6] [Medline: 32219180]
86. Flatt AA, Esco MR. Evaluating individual training adaptation with smartphone-derived heart rate variability in a collegiate female soccer team. J Strength Cond Res 2016;30(2):378-385. [doi: 10.1519/JSC.0000000000001095] [Medline: 26200192]
87. Frie K, Hartmann-Boyce J, Jebb S, Oke J, Aveyard P. Patterns in weight and physical activity tracking data preceding a stop in weight monitoring: observational analysis. J Med Internet Res 2020;22(3):e15790 [FREE Full text] [doi: 10.2196/15790] [Medline: 32181749]
88. Ghomrawi HM, Baumann LM, Kwon S, Hebal F, Hsiung G, Williams K, et al. Using accelerometers to characterize recovery after surgery in children. J Pediatr Surg 2018;53(8):1600-1605. [doi: 10.1016/j.jpedsurg.2017.09.016] [Medline: 29092769]
89. Girard MP, O'Shaughnessy J, Doucet C, Ruchat SM, Descarreaux M. Association between physical activity, weight loss, anxiety, and lumbopelvic pain in postpartum women. J Manipulative Physiol Ther 2020;43(6):655-666. [doi: 10.1016/j.jmpt.2019.11.008] [Medline: 32709518]
90. Gluck S, Summers MJ, Finnis ME, Andrawos A, Goddard TP, Hodgson CL, et al. An observational study investigating the use of patient-owned technology to quantify physical activity in survivors of critical illness. Aust Crit Care 2020;33(2):137-143. [doi: 10.1016/j.aucc.2019.01.009] [Medline: 30879879]
91. Green DR, Gerberich SG, Kim H, Ryan AD, McGovern PM, Church TR, et al. Janitor workload and occupational injuries. Am J Ind Med 2019;62(3):222-232. [doi: 10.1002/ajim.22940] [Medline: 30675912]
92. Gresham G, Hendifar AE, Spiegel B, Neeman E, Tuli R, Rimel BJ, et al. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. NPJ Digit Med 2018;1:27 [FREE Full text] [doi: 10.1038/s41746-018-0032-6] [Medline: 31304309]
93. Grimes L, Outtrim JG, Griffin SJ, Ercole A. Accelerometery as a measure of modifiable physical activity in high-risk elderly preoperative patients: a prospective observational pilot study. BMJ Open 2019;9(11):e032346 [FREE Full text] [doi: 10.1136/bmjopen-2019-032346] [Medline: 31685513]
94. Groat D, Kwon HJ, Grando MA, Cook CB, Thompson B. Comparing real-time self-tracking and device-recorded exercise data in subjects with type 1 diabetes. Appl Clin Inform 2018;9(4):919-926 [FREE Full text] [doi: 10.1055/s-0038-1676458] [Medline: 30586673]

XSL•FO
RenderX

95. Guo Y, Wang H, Zhang H, Chen Y, Lip GY. Population-based screening or targeted screening based on initial clinical risk assessment for atrial fibrillation: a report from the Huawei heart study. J Clin Med 2020;9(5):1493 [FREE Full text] [doi: 10.3390/jcm9051493] [Medline: 32429241]
96. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, MAFA II Investigators. Mobile photoplethysmographic technology to detect atrial fibrillation. J Am Coll Cardiol 2019;74(19):2365-2375. [doi: 10.1016/j.jacc.2019.08.019] [Medline: 31487545]
97. Hamed A, Curran C, Gwaltney C, DasMahapatra P. Mobility assessment using wearable technology in patients with late-onset Pompe disease. NPJ Digit Med 2019;2:70 [FREE Full text] [doi: 10.1038/s41746-019-0143-8] [Medline: 31341956]
98. Han M, Ye X, Preciado P, Williams S, Campos I, Bonner M, et al. Relationships between neighborhood walkability and objectively measured physical activity levels in hemodialysis patients. Blood Purif 2018;45(1-3):236-244 [FREE Full text] [doi: 10.1159/000485161] [Medline: 29478044]
99. Hemphill NM, Kuan MT, Harris KC. Reduced physical activity during COVID-19 pandemic in children with congenital heart disease. Can J Cardiol 2020;36(7):1130-1134 [FREE Full text] [doi: 10.1016/j.cjca.2020.04.038] [Medline: 32387502]
100. Hirschberg R, Sylvia LG, Wright EC, Gupta CT, McCarthy MD, Harward LK, et al. Collaborative songwriting intervention for veterans with post-traumatic stress disorder. J Altern Complement Med 2020;26(3):198-203. [doi: 10.1089/acm.2019.0368] [Medline: 31985263]
101. Huberty JL, Buman MP, Leiferman JA, Bushar J, Adams MA. Trajectories of objectively-measured physical activity and sedentary time over the course of pregnancy in women self-identified as inactive. Prev Med Rep 2016;3:353-360 [FREE Full text] [doi: 10.1016/j.pmedr.2016.04.004] [Medline: 27419036]
102. Jeong JN, Kim SH, Park KN. Relationship between objectively measured lifestyle factors and health factors in patients with knee osteoarthritis: the STROBE study. Medicine (Baltimore) 2019;98(26):e16060 [FREE Full text] [doi: 10.1097/MD.0000000000016060] [Medline: 31261513]
103. Jiang F, Kobayashi T, Ichihashi T, Ito K, Nomura S. The effect of citrus ginger aroma on heart rate and salivary cortisol level during sleep at home. IEEJ Trans Elec Electron Eng 2019;14(4):623-629. [doi: 10.1002/tee.22845]
104. John-Henderson NA, Palmer CA, Thomas A. Life stress, sense of belonging and sleep in American Indian college students. Sleep Health 2019;5(4):352-358 [FREE Full text] [doi: 10.1016/j.sleh.2019.04.001] [Medline: 31153800]
105. Jones MA, Skidmore PM, Stoner L, Harrex H, Saeedi P, Black K, et al. Associations of accelerometer-measured sedentary time, sedentary bouts, and physical activity with adiposity and fitness in children. J Sports Sci 2020;38(1):114-120. [doi: 10.1080/02640414.2019.1685842] [Medline: 31665975]
106. Jonker LT, Hendriks S, Lahr MM, van Munster BC, de Bock GH, van Leeuwen BL. Postoperative recovery of accelerometer-based physical activity in older cancer patients. Eur J Surg Oncol 2020;46(11):2083-2090 [FREE Full text] [doi: 10.1016/j.ejso.2020.06.012] [Medline: 32682650]
107. Kabbach EZ, Mazzuco A, Borghi-Silva A, Cabiddu R, Agnoleto AG, Barbosa JF, et al. Increased parasympathetic cardiac modulation in patients with acute exacerbation of COPD: how should we interpret it? Int J Chron Obstruct Pulmon Dis 2017;12:2221-2230 [FREE Full text] [doi: 10.2147/COPD.S134498] [Medline: 28814850]
108. Kagamiyama H, Yano R. Relationship between subjective fatigue, physical activity, and sleep indices in nurses working 16-hour night shifts in a rotating two-shift system. J Rural Med 2018;13(1):26-32 [FREE Full text] [doi: 10.2185/jrm.2951] [Medline: 29875894]
109. Kaura A, Sztriha L, Chan FK, Aeron-Thomas J, Gall N, Piechowski-Jozwiak B, et al. Early prolonged ambulatory cardiac monitoring in stroke (EPACS): an open-label randomised controlled trial. Eur J Med Res 2019;24(1):25 [FREE Full text] [doi: 10.1186/s40001-019-0383-8] [Medline: 31349792]
110. Kim DH, Nam KH, Choi BK, Han IH, Jeon TJ, Park SY. The usefulness of a wearable device in daily physical activity monitoring for the hospitalized patients undergoing lumbar surgery. J Korean Neurosurg Soc 2019;62(5):561-566 [FREE Full text] [doi: 10.3340/jkns.2018.0131] [Medline: 31337197]
111. Kim Y, Wijndaele K, Sharp SJ, Strain T, Pearce M, White T, et al. Specific physical activities, sedentary behaviours and sleep as long-term predictors of accelerometer-measured physical activity in 91,648 adults: a prospective cohort study. Int J Behav Nutr Phys Act 2019;16(1):41 [FREE Full text] [doi: 10.1186/s12966-019-0802-9] [Medline: 31064403]
112. Kimura N, Aso Y, Yabuuchi K, Ishibashi M, Hori D, Sasaki Y, et al. Association of modifiable lifestyle factors with cortical amyloid burden and cerebral glucose metabolism in older adults with mild cognitive impairment. JAMA Netw Open 2020;3(6):e205719 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.5719] [Medline: 32515796]
113. Kochiya Y, Hirabayashi A, Ichimaru Y. Nocturnal heart rate variability in 1-year-old infants analyzed by using the least square cosine spectrum method. J Physiol Anthropol 2017;36(1):36 [FREE Full text] [doi: 10.1186/s40101-017-0152-8] [Medline: 28915915]
114. Koehler F, Koehler K, Deckwart O, Prescher S, Wegscheider K, Kirwan BA, et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. Lancet 2018;392(10152):1047-1057 [FREE Full text] [doi: 10.1016/S0140-6736(18)31880-4] [Medline: 30153985]
115. Kolk D, Aarden JJ, MacNeil-Vroomen JL, Reichardt LA, van Seben R, van der Schaaf M, Hospital-ADL Study Group. Factors associated with step numbers in acutely hospitalized older adults: the hospital-activities of daily living study. J Am Med Dir Assoc 2021;22(2):425-432 [FREE Full text] [doi: 10.1016/j.jamda.2020.06.027] [Medline: 32713773]

XSL•FO
RenderX

116. Kruizinga MD, Zuiker RG, Sali E, de Kam ML, Doll RJ, Groeneveld GJ, et al. Finding suitable clinical endpoints for a potential treatment of a rare genetic disease: the case of ARID1B. Neurotherapeutics 2020;17(3):1300-1310 [FREE Full text] [doi: 10.1007/s13311-020-00868-9] [Medline: 32462407]
117. Lamar DL, Chou SH, Medverd JR, Swanson JO. Sedentary behavior in the workplace: a potential occupational hazard for radiologists. Curr Probl Diagn Radiol 2016;45(4):253-257. [doi: 10.1067/j.cpradiol.2015.10.007] [Medline: 26675263]
118. Larsen SC, Horgan G, Mikkelsen MK, Palmeira AL, Scott S, Duarte C, et al. Consistent sleep onset and maintenance of body weight after weight loss: an analysis of data from the NoHoW trial. PLoS Med 2020;17(7):e1003168 [FREE Full text] [doi: 10.1371/journal.pmed.1003168] [Medline: 32673309]
119. Lazaridou A, Paschali M, Schreiber K, Galenkamp L, Berry M, Paschalis T, et al. The association between daily physical exercise and pain among women with fibromyalgia: the moderating role of pain catastrophizing. Pain Rep 2020;5(4):e832 [FREE Full text] [doi: 10.1097/PR9.0000000000000832] [Medline: 32766468]
120. Le Hello C, Trombert B, Morel A, Chieh A, Brouard B, Boissier C. Performance analysis of walking of 10,000 regular users of a connected activity tracker. J Med Vasc 2018;43(4):231-237. [doi: 10.1016/j.jdmv.2018.04.001] [Medline: 29981731]
121. Lee JE, Lee DH, Oh TJ, Kim KM, Choi SH, Lim S, et al. Clinical feasibility of continuously monitored data for heart rate, physical activity, and sleeping by wearable activity trackers in patients with thyrotoxicosis: protocol for a prospective longitudinal observational study. JMIR Res Protoc 2018;7(2):e49 [FREE Full text] [doi: 10.2196/resprot.8119] [Medline: 29467121]
122. Lee JP, Chen S, Tsai CT, Chung HC, Chang WD. Characteristics associated with the differential activity of nondominant and dominant affected hands in patients with poststroke right hemiparesis. Occup Ther Int 2020;2020:2387378 [FREE Full text] [doi: 10.1155/2020/2387378] [Medline: 32565757]
123. Lee KM, Rogers MP, Galbarczyk A, Jasienska G, Clancy KB. Physical activity in women of reproductive age in a transitioning rural polish population. Am J Hum Biol 2019;31(3):e23231. [doi: 10.1002/ajhb.23231] [Medline: 30835924]
124. Liao Y, Robertson MC, Winne A, Wu IH, Le TA, Balachandran DD, et al. Investigating the within-person relationships between activity levels and sleep duration using Fitbit data. Transl Behav Med 2021;11(2):619-624. [doi: 10.1093/tbm/ibaa071] [Medline: 32667039]
125. Lim SE, Dodds R, Bacon D, Sayer AA, Roberts HC. Physical activity among hospitalised older people: insights from upper and lower limb accelerometry. Aging Clin Exp Res 2018;30(11):1363-1369 [FREE Full text] [doi: 10.1007/s40520-018-0930-0] [Medline: 29542070]
126. Lim WK, Davila S, Teo JX, Yang C, Pua CJ, Blöcker C, et al. Beyond fitness tracking: the use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. PLoS Biol 2018;16(2):e2004285 [FREE Full text] [doi: 10.1371/journal.pbio.2004285] [Medline: 29485983]
127. Lin H, Sardana M, Zhang Y, Liu C, Trinquart L, Benjamin EJ, et al. Association of habitual physical activity with cardiovascular disease risk. Circ Res 2020;127(10):1253-1260. [doi: 10.1161/CIRCRESAHA.120.317578] [Medline: 32842915]
128. Liu S, Eaton CB, Driban JB, McAlindon TE, Lapane KL. Comparison of self-report and objective measures of physical activity in US adults with osteoarthritis. Rheumatol Int 2016;36(10):1355-1364. [doi: 10.1007/s00296-016-3537-9] [Medline: 27435920]
129. Low CA, Bovbjerg DH, Ahrendt S, Choudry MH, Holtzman M, Jones HL, et al. Fitbit step counts during inpatient recovery from cancer surgery as a predictor of readmission. Ann Behav Med 2018;52(1):88-92 [FREE Full text] [doi: 10.1093/abm/kax022] [Medline: 29538623]
130. Low CA, Dey AK, Ferreira D, Kamarck T, Sun W, Bae S, et al. Estimation of symptom severity during chemotherapy from passively sensed data: exploratory study. J Med Internet Res 2017;19(12):e420 [FREE Full text] [doi: 10.2196/jmir.9046] [Medline: 29258977]
131. Ma J, Martin R, Chan B, Gofeld M, Geary MP, Laffey JG, et al. Using activity trackers to quantify postpartum ambulation: a prospective observational study of ambulation after regional anesthesia and analgesia interventions. Anesthesiology 2018;128(3):598-608 [FREE Full text] [doi: 10.1097/ALN.0000000000001979] [Medline: 29135475]
132. Maijala A, Kinnunen H, Koskimäki H, Jämsä T, Kangas M. Nocturnal finger skin temperature in menstrual cycle tracking: ambulatory pilot study using a wearable Oura ring. BMC Womens Health 2019;19(1):150 [FREE Full text] [doi: 10.1186/s12905-019-0844-9] [Medline: 31783840]
133. Makic MB, Gilbert D, Jankowski C, Reeder B, Al-Salmi N, Starr W, et al. Sensor and survey measures associated with daily fatigue in HIV: findings from a mixed-method study. J Assoc Nurses AIDS Care 2020;31(1):12-24. [doi: 10.1097/JNC.0000000000000152] [Medline: 31860594]
134. Massey WV, Stellino MB, Geldhof J. An observational study of recess quality and physical activity in urban primary schools. BMC Public Health 2020;20(1):792 [FREE Full text] [doi: 10.1186/s12889-020-08849-5] [Medline: 32460860]
135. Massey WV, Stellino MB, Fraser M. Individual and environmental correlates of school-based recess engagement. Prev Med Rep 2018;11:247-253 [FREE Full text] [doi: 10.1016/j.pmedr.2018.07.005] [Medline: 30094127]

XSL•FO
RenderX

136. Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, RADAR-CNS consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry 2019;19(1):72 [FREE Full text] [doi: 10.1186/s12888-019-2049-z] [Medline: 30777041]
137. McLean DC, Nakamura J, Csikszentmihalyi M. Reconsidering the experience machine: self-reported versus objective measures of physical activity to increase positive affect. J Health Psychol 2020;25(13-14):2428-2439. [doi: 10.1177/1359105318801939] [Medline: 30246550]
138. Mead MP, Baron K, Sorby M, Irish LA. Daily associations between sleep and physical activity. Int J Behav Med 2019;26(5):562-568. [doi: 10.1007/s12529-019-09810-6] [Medline: 31372835]
139. Meijer K, Annegarn J, Lima Passos V, Savelberg HH, Schols AM, Wouters EF, et al. Characteristics of daily arm activities in patients with COPD. Eur Respir J 2014;43(6):1631-1641 [FREE Full text] [doi: 10.1183/09031936.00082513] [Medline: 24558175]
140. Melin M, Hagerman I, Gonon A, Gustafsson T, Rullman E. Variability in physical activity assessed with accelerometer is an independent predictor of mortality in CHF patients. PLoS One 2016;11(4):e0153036 [FREE Full text] [doi: 10.1371/journal.pone.0153036] [Medline: 27054323]
141. Menai M, Brouard B, Vegreville M, Chieh A, Schmidt N, Oppert JM, et al. Cross-sectional and longitudinal associations of objectively-measured physical activity on blood pressure: evaluation in 37 countries. Health Promot Perspect 2017;7(4):190-196 [FREE Full text] [doi: 10.15171/hpp.2017.34] [Medline: 29085795]
142. Mendelsohn D, Despot I, Gooderham PA, Singhal A, Redekop GJ, Toyota BD. Impact of work hours and sleep on well-being and burnout for physicians-in-training: the Resident Activity Tracker Evaluation Study. Med Educ 2019;53(3):306-315. [doi: 10.1111/medu.13757] [Medline: 30485496]
143. Mičková E, Machová K, Daďová K, Svobodová I. Does dog ownership affect physical activity, sleep, and self-reported health in older adults? Int J Environ Res Public Health 2019;16(18):3355 [FREE Full text] [doi: 10.3390/ijerph16183355] [Medline: 31514379]
144. Miller DJ, Capodilupo JV, Lastella M, Sargent C, Roach GD, Lee VH, et al. Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. PLoS One 2020;15(12):e0243693 [FREE Full text] [doi: 10.1371/journal.pone.0243693] [Medline: 33301493]
145. Mobbs RJ, Phan K, Maharaj M, Rao PJ. Physical activity measured with accelerometer and self-rated disability in lumbar spine surgery: a prospective study. Global Spine J 2016;6(5):459-464 [FREE Full text] [doi: 10.1055/s-0035-1565259] [Medline: 27433430]
146. Mocny-Pachońska K, Doniec R, Trzcionka A, Pachoński M, Piaseczna N, Sieciński S, et al. Evaluating the stress-response of dental students to the dental school environment. PeerJ 2020;8:e8981 [FREE Full text] [doi: 10.7717/peerj.8981] [Medline: 32411517]
147. Modarress-Sadeghi M, Oura P, Junno JA, Niemelä M, Niinimäki J, Jämsä T, et al. Objectively measured physical activity is associated with vertebral size in midlife. Med Sci Sports Exerc 2019;51(8):1606-1612. [doi: 10.1249/MSS.0000000000001962] [Medline: 30817715]
148. Mora-Gonzalez J, Migueles JH, Esteban-Cornejo I, Cadenas-Sanchez C, Pastor-Villaescusa B, Molina-García P, et al. Sedentarism, physical activity, steps, and neurotrophic factors in obese children. Med Sci Sports Exerc 2019;51(11):2325-2333. [doi: 10.1249/MSS.0000000000002064] [Medline: 31634295]
149. Morcos MW, Teeter MG, Somerville LE, Lanting B. Correlation between hip osteoarthritis and the level of physical activity as measured by wearable technology and patient-reported questionnaires. J Orthop 2019;20:236-239 [FREE Full text] [doi: 10.1016/j.jor.2019.11.049] [Medline: 32071522]
150. Morhardt DR, Luckenbaugh A, Goldstein C, Faerber GJ. Determining resident sleep during and after call with commercial sleep monitoring devices. Urology 2017;106:39-44. [doi: 10.1016/j.urology.2017.03.059] [Medline: 28502597]
151. Mueller C, Stollfuss B, Roitenberg A, Harder J, Richter MJ. Evaluation of clinical outcomes and simultaneous digital tracking of daily physical activity, heart rate, and inhalation behavior in patients with pulmonary arterial hypertension treated with inhaled iloprost: protocol for the observational VENTASTEP study. JMIR Res Protoc 2019;8(4):e12144 [FREE Full text] [doi: 10.2196/12144] [Medline: 30985279]
152. Murphree DH, Kinard TN, Khera N, Storlie CB, Ngufor C, Upadhyaya S, et al. Measuring the impact of ambulatory red blood cell transfusion on home functional status: study protocol for a pilot randomized controlled trial. Trials 2017;18(1):153 [FREE Full text] [doi: 10.1186/s13063-017-1873-z] [Medline: 28359342]
153. Natarajan A, Pantelopoulos A, Emir-Farinas H, Natarajan P. Heart rate variability with photoplethysmography in 8 million individuals: a cross-sectional study. Lancet Digit Health 2020;2(12):e650-e657. [doi: 10.1016/S2589-7500(20)30246-6] [Medline: 33328029]
154. Niemelä MS, Kangas M, Ahola RJ, Auvinen JP, Leinonen AM, Tammelin TH, et al. Dose-response relation of self-reported and accelerometer-measured physical activity to perceived health in middle age-the Northern Finland Birth Cohort 1966 Study. BMC Public Health 2019;19(1):21 [FREE Full text] [doi: 10.1186/s12889-018-6359-8] [Medline: 30612541]
155. Nishida M, Kikuchi S, Miwakeichi F, Suda S. Night duty and decreased brain activity of medical residents: a wearable optical topography study. Med Educ Online 2017;22(1):1379345 [FREE Full text] [doi: 10.1080/10872981.2017.1379345] [Medline: 28954586]

XSL•FO
RenderX

156. Nowell WB, Curtis JR, Nolot SK, Curtis D, Venkatachalam S, Owensby JK, et al. Digital tracking of rheumatoid arthritis longitudinally (DIGITAL) using biosensor and patient-reported outcome data: protocol for a real-world study. JMIR Res Protoc 2019;8(9):e14665 [FREE Full text] [doi: 10.2196/14665] [Medline: 31573949]
157. Orme MW, Steiner MC, Morgan MD, Kingsnorth AP, Esliger DW, Singh SJ, et al. 24-hour accelerometry in COPD: exploring physical activity, sedentary behavior, sleep and clinical characteristics. Int J Chron Obstruct Pulmon Dis 2019;14:419-430 [FREE Full text] [doi: 10.2147/COPD.S183029] [Medline: 30863042]
158. Otsuki T, Ishii N. Association between blood pressure changes during self-paced outdoor walking and air temperature. Clin Physiol Funct Imaging 2017;37(2):155-161. [doi: 10.1111/cpf.12280] [Medline: 26211639]
159. Pakhomov SV, Thuras PD, Finzel R, Eppel J, Kotlyar M. Using consumer-wearable technology for remote assessment of physiological response to stress in the naturalistic environment. PLoS One 2020;15(3):e0229942 [FREE Full text] [doi: 10.1371/journal.pone.0229942] [Medline: 32210441]
160. Pastor N, Khalilian E, Caballeria E, Morrison D, Sanchez Luque U, Matrai S, et al. Remote monitoring telemedicine (REMOTE) platform for patients with anxiety symptoms and alcohol use disorder: protocol for a case-control study. JMIR Res Protoc 2020;9(6):e16964 [FREE Full text] [doi: 10.2196/16964] [Medline: 32579124]
161. Peach HD, Gaultney JF, Ruggiero AR. Direct and indirect associations of sleep knowledge and attitudes with objective and subjective sleep duration and quality via sleep hygiene. J Prim Prev 2018;39(6):555-570. [doi: 10.1007/s10935-018-0526-7] [Medline: 30443688]
162. Pépin JL, Bruno RM, Yang RY, Vercamer V, Jouhaud P, Escourrou P, et al. Wearable activity trackers for monitoring adherence to home confinement during the COVID-19 pandemic worldwide: data aggregation and analysis. J Med Internet Res 2020;22(6):e19787 [FREE Full text] [doi: 10.2196/19787] [Medline: 32501803]
163. Phillips LJ, Petroski GF, Conn VS, Brown M, Leary E, Teri L, et al. Exploring path models of disablement in residential care and assisted living residents. J Appl Gerontol 2018;37(12):1490-1516 [FREE Full text] [doi: 10.1177/0733464816672048] [Medline: 27708073]
164. Pozehl BJ, Mcguire R, Duncan K, Hertzog M, Deka P, Norman J, et al. Accelerometer-measured daily activity levels and related factors in patients with heart failure. J Cardiovasc Nurs 2018;33(4):329-335 [FREE Full text] [doi: 10.1097/JCN.0000000000000464] [Medline: 29538050]
165. Pradhan S, Kelly VE. Quantifying physical activity in early Parkinson disease using a commercial activity monitor. Parkinsonism Relat Disord 2019;66:171-175 [FREE Full text] [doi: 10.1016/j.parkreldis.2019.08.001] [Medline: 31420310]
166. Quer G, Gouda P, Galarnyk M, Topol EJ, Steinhubl SR. Inter- and intraindividual variability in daily resting heart rate and its associations with age, sex, sleep, BMI, and time of year: retrospective, longitudinal cohort study of 92,457 adults. PLoS One 2020;15(2):e0227709 [FREE Full text] [doi: 10.1371/journal.pone.0227709] [Medline: 32023264]
167. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. Nat Med 2021;27(1):73-77. [doi: 10.1038/s41591-020-1123-x] [Medline: 33122860]
168. Quiroz JC, Geangu E, Yong MH. Emotion recognition using smart watch sensor data: mixed-design study. JMIR Ment Health 2018;5(3):e10153 [FREE Full text] [doi: 10.2196/10153] [Medline: 30089610]
169. Rahman MJ, Morshed BI, Harmon B. A field study to capture events of interest (EoI) from living labs using wearables for spatiotemporal monitoring towards a framework of smart health (sHealth). Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5943-5947. [doi: 10.1109/EMBC44109.2020.9175771] [Medline: 33019327]
170. Rahman MJ, Nemati E, Rahman M, Vatanparvar K, Nathan V, Kuang J. Toward early severity assessment of obstructive lung disease using multi-modal wearable sensor data fusion during walking. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5935-5938. [doi: 10.1109/EMBC44109.2020.9176559] [Medline: 33019325]
171. Ramirez E, Marinsek N, Bradshaw B, Kanard R, Foschini L. Continuous digital assessment for weight loss surgery patients. Digit Biomark 2020;4(1):13-20 [FREE Full text] [doi: 10.1159/000506417] [Medline: 32399512]
172. Raywood E, Douglas H, Kapoor K, Filipow N, Murray N, O'Connor R, et al. Protocol for Project Fizzyo, an analytic longitudinal observational cohort study of physiotherapy for children and young people with cystic fibrosis, with interrupted time-series design. BMJ Open 2020;10(10):e039587 [FREE Full text] [doi: 10.1136/bmjopen-2020-039587] [Medline: 33033031]
173. Reddy R, El Youssef J, Winters-Stone K, Branigan D, Leitschuh J, Castle J, et al. The effect of exercise on sleep in adults with type 1 diabetes. Diabetes Obes Metab 2018;20(2):443-447 [FREE Full text] [doi: 10.1111/dom.13065] [Medline: 28718987]
174. Roe J, Salmon L, Twiggs J. Objective measure of activity level after total knee arthoplasty with the use of the 'Fitbit' device. Orthop J Sports Med 2016;4(Supplement 2):2325967116S00012 [FREE Full text] [doi: 10.1177/2325967116S00012]
175. Root ED, Graney B, Baird S, Churney T, Fier K, Korn M, et al. Physical activity and activity space in patients with pulmonary fibrosis not prescribed supplemental oxygen. BMC Pulm Med 2017;17(1):154 [FREE Full text] [doi: 10.1186/s12890-017-0495-2] [Medline: 29169394]
176. Rykov Y, Thach TQ, Dunleavy G, Roberts AC, Christopoulos G, Soh CK, et al. Activity tracker-based metrics as digital markers of cardiometabolic health in working adults: cross-sectional study. JMIR Mhealth Uhealth 2020;8(1):e16409 [FREE Full text] [doi: 10.2196/16409] [Medline: 32012098]

XSL•FO
RenderX

177. Saarikko J, Niela-Vilen H, Ekholm E, Hamari L, Azimi I, Liljeberg P, et al. Continuous 7-month internet of things-based monitoring of health parameters of pregnant and postpartum women: prospective observational feasibility study. JMIR Form Res 2020;4(7):e12417 [FREE Full text] [doi: 10.2196/12417] [Medline: 32706696]
178. Saif N, Yan P, Niotis K, Scheyer O, Rahman A, Berkowitz M, et al. Feasibility of using a wearable biosensor device in patients at risk for Alzheimer's disease dementia. J Prev Alzheimers Dis 2020;7(2):104-111 [FREE Full text] [doi: 10.14283/jpad.2019.39] [Medline: 32236399]
179. Scarlett S, Nolan H, Kenny RA, O'Connell MD. Objective sleep duration in older adults: results from the Irish longitudinal study on ageing. J Am Geriatr Soc 2020;68(1):120-128. [doi: 10.1111/jgs.16177] [Medline: 31579942]
180. Sehgal S, Chowdhury A, Rabih F, Gadre A, Park MM, Li M, STep-count using an Accelerometer for Monitoring PAH—STAMP Study group. Counting steps: a new way to monitor patients with pulmonary arterial hypertension. Lung 2019;197(4):501-508. [doi: 10.1007/s00408-019-00239-y] [Medline: 31144016]
181. Semaan S, Dewland TA, Tison GH, Nah G, Vittinghoff E, Pletcher MJ, et al. Physical activity and atrial fibrillation: data from wearable fitness trackers. Heart Rhythm 2020;17(5 Pt B):842-846. [doi: 10.1016/j.hrthm.2020.02.013] [Medline: 32354448]
182. Shaw RJ, Barnes A, Steinberg D, Vaughn J, Diane A, Levine E, et al. Enhancing diabetes self-management through collection and visualization of data from multiple mobile health technologies: protocol for a development and feasibility trial. JMIR Res Protoc 2019;8(6):e13517 [FREE Full text] [doi: 10.2196/13517] [Medline: 31162127]
183. Shen H, Zhao J, Zhou X, Li J, Wan Q, Huang J, et al. Impaired chronotropic response to physical activities in heart failure patients. BMC Cardiovasc Disord 2017;17(1):136 [FREE Full text] [doi: 10.1186/s12872-017-0571-9] [Medline: 28545575]
184. Shilaih M, Goodale BM, Falco L, Kübler F, De Clerck V, Leeners B. Modern fertility awareness methods: wrist wearables capture the changes in temperature associated with the menstrual cycle. Biosci Rep 2018;38(6):BSR20171279 [FREE Full text] [doi: 10.1042/BSR20171279] [Medline: 29175999]
185. Shilaih M, Clerck V, Falco L, Kübler F, Leeners B. Pulse rate measurement during sleep using wearable sensors, and its correlation with the menstrual cycle phases, a prospective observational study. Sci Rep 2017;7(1):1294 [FREE Full text] [doi: 10.1038/s41598-017-01433-9] [Medline: 28465583]
186. Shufelt CL, Kim A, Joung S, Barsky L, Arnold C, Cheng S, et al. Biometric and psychometric remote monitoring and cardiovascular risk biomarkers in ischemic heart disease. J Am Heart Assoc 2020;9(18):e016023 [FREE Full text] [doi: 10.1161/JAHA.120.016023] [Medline: 32896202]
187. Silverman-Lloyd LG, Kianoush S, Blaha MJ, Sabina AB, Graham GN, Martin SS. mActive-Smoke: a prospective observational study using mobile health tools to assess the association of physical activity with smoking urges. JMIR Mhealth Uhealth 2018;6(5):e121 [FREE Full text] [doi: 10.2196/mhealth.9292] [Medline: 29752250]
188. Smith SL, Ludy MJ, Tucker RM. Changes in taste preference and steps taken after sleep curtailment. Physiol Behav 2016;163:228-233. [doi: 10.1016/j.physbeh.2016.05.002] [Medline: 27184237]
189. Solomon MD, Yang J, Sung SH, Livingston ML, Sarlas G, Lenane JC, et al. Incidence and timing of potentially high-risk arrhythmias detected through long term continuous ambulatory electrocardiographic monitoring. BMC Cardiovasc Disord 2016;16:35 [FREE Full text] [doi: 10.1186/s12872-016-0210-x] [Medline: 26883019]
190. Souza RT, Cecatti JG, Mayrink J, Galvão RB, Costa ML, Feitosa F, MAES-I Study Group. Identification of earlier predictors of pregnancy complications through wearable technologies in a Brazilian multicentre cohort: Maternal Actigraphy Exploratory Study I (MAES-I) study protocol. BMJ Open 2019;9(4):e023101 [FREE Full text] [doi: 10.1136/bmjopen-2018-023101] [Medline: 31005906]
191. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA 2018;320(2):146-155 [FREE Full text] [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
192. Stelzer EM, Book S, Graessel E, Hofner B, Kornhuber J, Luttenberger K. Bouldering psychotherapy reduces depressive symptoms even when general physical activity is controlled for: a randomized controlled trial. Heliyon 2018;4(3):e00580 [FREE Full text] [doi: 10.1016/j.heliyon.2018.e00580] [Medline: 29862344]
193. Strain T, Wijndaele K, Dempsey PC, Sharp SJ, Pearce M, Jeon J, et al. Wearable-device-measured physical activity and future health risk. Nat Med 2020;26(9):1385-1391. [doi: 10.1038/s41591-020-1012-3] [Medline: 32807930]
194. Tateishi K, Ohtani N, Ohta M. Physiological effects of interactions between female dog owners with neuroticism and their dogs. J Vet Behav 2014;9(6):304-310. [doi: 10.1016/j.jveb.2014.08.005]
195. Thijs I, Fresiello L, Oosterlinck W, Sinnaeve P, Rega F. Assessment of physical activity by wearable technology during rehabilitation after cardiac surgery: explorative prospective monocentric observational cohort study. JMIR Mhealth Uhealth 2019;7(1):e9865 [FREE Full text] [doi: 10.2196/mhealth.9865] [Medline: 30702433]
196. Thomas S, Yingling L, Adu-Brimpong J, Mitchell V, Ayers CR, Wallen GR, et al. Mobile health technology can objectively capture physical activity (PA) targets among African-American women within resource-limited communities-the Washington, D.C. cardiovascular health and needs assessment. J Racial Ethn Health Disparities (forthcoming) 2016:- (forthcoming) [FREE Full text] [doi: 10.1007/s40615-016-0290-4] [Medline: 27913983]
197. Thompson JF, Severson RL, Rosecrance JC. Occupational physical activity in brewery and office workers. J Occup Environ Hyg 2018;15(9):686-699. [doi: 10.1080/15459624.2018.1492136] [Medline: 30188781]

XSL•FO
RenderX

198. Thorup C, Hansen J, Grønkjær M, Andreasen JJ, Nielsen G, Sørensen EE, et al. Cardiac patients' walking activity determined by a step counter in cardiac telerehabilitation: data from the intervention arm of a randomized controlled trial. J Med Internet Res 2016;18(4):e69 [FREE Full text] [doi: 10.2196/jmir.5191] [Medline: 27044310]
199. Tomitani N, Kanegae H, Suzuki Y, Kuwabara M, Kario K. Stress-induced blood pressure elevation self-measured by a wearable watch-type device. Am J Hypertens 2021;34(4):377-382 [FREE Full text] [doi: 10.1093/ajh/hpaa139] [Medline: 32852527]
200. Turakhia MP, Ullal AJ, Hoang DD, Than CT, Miller JD, Friday KJ, et al. Feasibility of extended ambulatory electrocardiogram monitoring to identify silent atrial fibrillation in high-risk patients: the Screening Study for Undiagnosed Atrial Fibrillation (STUDY-AF). Clin Cardiol 2015;38(5):285-292 [FREE Full text] [doi: 10.1002/clc.22387] [Medline: 25873476]
201. Turel O, Romashkin A, Morrison KM. Health outcomes of information system use lifestyles among adolescents: videogame addiction, sleep curtailment and cardio-metabolic deficiencies. PLoS One 2016;11(5):e0154764 [FREE Full text] [doi: 10.1371/journal.pone.0154764] [Medline: 27149512]
202. Twiggs J, Salmon L, Kolos E, Bogue E, Miles B, Roe J. Measurement of physical activity in the pre- and early post-operative period after total knee arthroplasty for osteoarthritis using a Fitbit Flex device. Med Eng Phys 2018;51:31-40. [doi: 10.1016/j.medengphy.2017.10.007] [Medline: 29117912]
203. van der Kamp MR, Thio BJ, Tabak M, Hermens HJ, Driessen J, van der Palen J. Does exercise-induced bronchoconstriction affect physical activity patterns in asthmatic children? J Child Health Care 2020;24(4):577-588. [doi: 10.1177/1367493519881257] [Medline: 31607144]
204. van Wamelen DJ, Hota S, Podlewska A, Leta V, Trivedi D, Rizos A, et al. Non-motor correlates of wrist-worn wearable sensor use in Parkinson's disease: an exploratory analysis. NPJ Parkinsons Dis 2019;5:22 [FREE Full text] [doi: 10.1038/s41531-019-0094-4] [Medline: 31602393]
205. Vaughn NH, Dunklebarger MF, Mason MW. Individual patient-reported activity levels before and after joint arthroplasty are neither accurate nor reproducible. Clin Orthop Relat Res 2019;477(3):536-544 [FREE Full text] [doi: 10.1097/CORR.0000000000000591] [Medline: 30543533]
206. Venkataramanan R, Thirunarayan K, Jaimini U, Kadariya D, Yip HY, Kalra M, et al. Determination of personalized asthma triggers from multimodal sensing and a mobile app: observational study. JMIR Pediatr Parent 2019;2(1):e14300 [FREE Full text] [doi: 10.2196/14300] [Medline: 31518318]
207. Wallace M, Downing N, Lourens S, Mills J, Kim JI, Long J, Predict-Hd Investigators And Coordinators Of The Huntington Study Group. Is there an association of physical activity with brain volume, behavior, and day-to-day functioning? A cross sectional design in prodromal and early Huntington disease. PLoS Curr 2016;8:ecurrents.hd.cba6ea74972cf8412a73ce52eb018c1e [FREE Full text] [doi: 10.1371/currents.hd.cba6ea74972cf8412a73ce52eb018c1e] [Medline: 27818843]
208. Wang C, Lizardo O, Hachen DS. Using Fitbit data to monitor the heart rate evolution patterns of college students. J Am Coll Health 2020:1-8. [doi: 10.1080/07448481.2020.1775610] [Medline: 32569509]
209. Weatherall J, Paprocki Y, Meyer TM, Kudel I, Witt EA. Sleep tracking and exercise in patients with type 2 diabetes mellitus (Step-D): pilot study to determine correlations between Fitbit data and patient-reported outcomes. JMIR Mhealth Uhealth 2018;6(6):e131 [FREE Full text] [doi: 10.2196/mhealth.8122] [Medline: 29871856]
210. Weeks DL, Sprint GL, Stilwill V, Meisen-Vehrs AL, Cook DJ. Implementing wearable sensors for continuous assessment of daytime heart rate response in inpatient rehabilitation. Telemed J E Health 2018;24(12):1014-1020 [FREE Full text] [doi: 10.1089/tmj.2017.0306] [Medline: 29608421]
211. Wiles T, Kellogg D, Furr H, Nessler JA, Newcomer SC. Characterization of adult heart rate responses during recreational skateboarding at community skateparks. Int J Exerc Sci 2020;13(2):501-510 [FREE Full text] [Medline: 32509111]
212. Winnebeck EC, Fischer D, Leise T, Roenneberg T. Dynamics and ultradian structure of human sleep in real life. Curr Biol 2018;28(1):49-59.e5 [FREE Full text] [doi: 10.1016/j.cub.2017.11.063] [Medline: 29290561]
213. Wyatt KD, Poole LR, Mullan AF, Kopecky SL, Heaton HA. Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using Apple watch. J Am Med Inform Assoc 2020;27(9):1359-1363 [FREE Full text] [doi: 10.1093/jamia/ocaa137] [Medline: 32979046]
214. Xu X, Conomos MP, Manor O, Rohwer JE, Magis AT, Lovejoy JC. Habitual sleep duration and sleep duration variation are independently associated with body mass index. Int J Obes (Lond) 2018;42(4):794-800. [doi: 10.1038/ijo.2017.223] [Medline: 28895585]
215. Yang Y, Hirdes JP, Dubin JA, Lee J. Fall risk classification in community-dwelling older adults using a smart wrist-worn device and the resident assessment instrument-home care: prospective observational study. JMIR Aging 2019;2(1):e12153 [FREE Full text] [doi: 10.2196/12153] [Medline: 31518278]
216. Youn IH, Lee JM. Seafarers' physical activity and sleep patterns: results from Asia-Pacific sea routes. Int J Environ Res Public Health 2020;17(19):7266 [FREE Full text] [doi: 10.3390/ijerph17197266] [Medline: 33027892]
217. Yu Y, Chen Z, Bu J, Zhang Q. Do stairs inhibit seniors who live on upper floors from going out? HERD 2020;13(4):128-143. [doi: 10.1177/1937586720936588] [Medline: 32666833]

XSL•FO
RenderX

218. Zhu G, Catt M, Cassidy S, Birch-Machin M, Trenell M, Hiden H, et al. Objective sleep assessment in >80,000 UK mid-life adults: associations with sociodemographic characteristics, physical activity and caffeine. PLoS One 2019;14(12):e0226220 [FREE Full text] [doi: 10.1371/journal.pone.0226220] [Medline: 31881028]
219. Zhuo K, Gao C, Wang X, Zhang C, Wang Z. Stress and sleep: a survey based on wearable sleep trackers among medical and nursing staff in Wuhan during the COVID-19 pandemic. Gen Psychiatr 2020;33(3):e100260 [FREE Full text] [doi: 10.1136/gpsych-2020-100260] [Medline: 32596641]
220. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021;10(1):89 [FREE Full text] [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]
221. Create your own custom map. MapChart. URL: https://mapchart.net/ [accessed 2022-01-19]
222. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
223. Gandomi A, Haider M. Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manage 2015;35(2):137-144. [doi: 10.1016/j.ijinfomgt.2014.10.007]
224. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. Am J Physiol Regul Integr Comp Physiol 2017;312(3):R358-R367. [doi: 10.1152/ajpregu.00349.2016] [Medline: 28052867]
225. Kristoffersson A, Lindén M. A systematic review on the use of wearable body sensors for health monitoring: a qualitative synthesis. Sensors (Basel) 2020;20(5):15902 [FREE Full text] [doi: 10.3390/s20051502] [Medline: 32182907]
226. Nelson BW, Low CA, Jacobson N, Areán P, Torous J, Allen NB. Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research. NPJ Digit Med 2020;3:90 [FREE Full text] [doi: 10.1038/s41746-020-0297-4] [Medline: 32613085]
227. Who are the world's biggest wearable tech buyers? Ipsos. 2018. URL: https://www.ipsos.com/en-ro/who-are-worlds-biggest-wearable-tech-buyers [accessed 2021-05-20]
228. Vailshery LS. Number of connected wearable devices worldwide by region from 2015 to 2022. Statista. 2021. URL: https://www.statista.com/statistics/490231/wearable-devices-worldwide-by-region/ [accessed 2021-05-20]
229. What is GDPR, the EU's new data protection law? General Data Protection Regulation.: European Union; 2018. URL: https://gdpr.eu/what-is-gdpr/ [accessed 2021-06-08]
230. Blankertz A, Specht L. On regulation for data trusts: policy brief. Stiftung Neue Verantwortung. 2021. URL: https://www.stiftung-nv.de/en/publication/regulation-data-trusts [accessed 2021-07-06]
231. Populations underrepresented in the extramural scientific workforce. Scientific Workforce Diversity.: National Institutes of Health; 2020. URL: https://diversity.nih.gov/about-us/population-underrepresented [accessed 2021-10-13]
232. Evans GF, Shirk A, Muturi P, Soliman EZ. Feasibility of using mobile ECG recording technology to detect atrial fibrillation in low-resource settings. Glob Heart 2017;12(4):285-289. [doi: 10.1016/j.gheart.2016.12.003] [Medline: 28302547]
233. Hicks JL, Althoff T, Sosic R, Kuhar P, Bostjancic B, King AC, et al. Best practices for analyzing large-scale health data from wearables and smartphone apps. NPJ Digit Med 2019;2:45 [FREE Full text] [doi: 10.1038/s41746-019-0121-1] [Medline: 31304391]
234. Xu S, Rwei AY, Vwalika B, Chisembele MP, Stringer JS, Ginsburg AS, et al. Wireless skin sensors for physiological monitoring of infants in low-income and middle-income countries. Lancet Digit Health 2021;3(4):e266-e273 [FREE Full text] [doi: 10.1016/S2589-7500(21)00001-7] [Medline: 33640306]
235. Sauerborn R. A gaping research gap regarding the climate change impact on health in poor countries. Eur J Epidemiol 2017;32(9):855-856 [FREE Full text] [doi: 10.1007/s10654-017-0258-7] [Medline: 28573342]
236. Barteit S, Boudo V, Ouedraogo A, Zabré P, Ouremi L, Sié A, et al. Feasibility, acceptability and validation of wearable devices for climate change and health research in the low-resource contexts of Burkina Faso and Kenya: study protocol. PLoS One 2021;16(9):e0257170 [FREE Full text] [doi: 10.1371/journal.pone.0257170] [Medline: 34591893]
237. Lam CK, Hang J, Zhang D, Wang Q, Ren M, Huang C. Effects of short-term physiological and psychological adaptation on summer thermal comfort of outdoor exercising people in China. Build Environ 2021;198:107877. [doi: 10.1016/j.buildenv.2021.107877]

Abbreviations

AF: atrial fibrillation
API: application programing interface
CE: Communauté Européenne
ECG: electrocardiogram
FDA: Food and Drug Administration
HR: heart rate
PA: physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

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

Edited by A Mavragani; submitted 20.10.21; peer-reviewed by H Zhang, P Escourrou, A Henriksen; comments to author 12.11.21; revised version received 23.11.21; accepted 17.12.21; published 25.01.22.

Please cite as:
Huhn S, Axt M, Gunga HC, Maggioni MA, Munga S, Obor D, Sié A, Boudo V, Bunker A, Sauerborn R, Bärnighausen T, Barteit S
The Impact of Wearable Technologies in Health Research: Scoping Review
JMIR Mhealth Uhealth 2022;10(1):e34384
URL: https://mhealth.jmir.org/2022/1/e34384
doi:10.2196/34384
PMID:35076409

©Sophie Huhn, Miriam Axt, Hanns-Christian Gunga, Martina Anna Maggioni, Stephen Munga, David Obor, Ali Sié, Valentin Boudo, Aditi Bunker, Rainer Sauerborn, Till Bärnighausen, Sandra Barteit. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 25.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

The Influence of Design and Implementation Characteristics on the Use of Maternal Mobile Health Interventions in Kenya: Systematic Literature Review

Karen Sowon[1], BBIT, MScIT, PhD; Priscilla Maliwichi[1,2], BSc, MSc; Wallace Chigona[1], BSocSci, MSci, PhD

[1]University of Cape Town, Department of Information Systems, Cape Town, South Africa

[2]Department of Computer Science and Information Technology, Malawi University of Science and Technology, Limbe, Malawi

Corresponding Author:
Karen Sowon, BBIT, MScIT, PhD
University of Cape Town
Department of Information Systems
University of Cape Town
Cape Town, 7700
South Africa
Phone: 27 651932568
Email: swnkar004@myuct.ac.za

Abstract

Background: The growth of mobile technology in developing countries, coupled with pressing maternal health care challenges, has led to a widespread implementation of maternal mobile health (mHealth) innovations. However, reviews generating insights on how the characteristics of the interventions influence use are scarce.

Objective: This study aims to review maternal mHealth interventions in Kenya to explore the influence of intervention design and implementation characteristics on use by maternal health clients. We also provide a starting inventory for maternal mHealth interventions in the country.

Methods: Using a systematic approach, we retrieved a total of 1100 citations from both peer-reviewed and gray sources. Articles were screened on the basis of an inclusion and exclusion criterion, and the results synthesized by categorizing and characterizing the interventions presented in the articles. The first phase of the literature search was conducted between January and April 2019, and the second phase was conducted between April and June 2021.

Results: A total of 16 articles were retrieved, comprising 13 maternal mHealth interventions. The study highlighted various mHealth design and implementation characteristics that may influence the use of these interventions.

Conclusions: In addition to elaborating on insights that would be useful in the design and implementation of future interventions, this study contributes to a local inventory of maternal mHealth interventions that may be useful to researchers and implementers in mHealth. This study highlights the need for explanatory studies to elucidate maternal mHealth use, while complementing existing evidence on mHealth effectiveness.

(JMIR Mhealth Uhealth 2022;10(1):e22093) doi:10.2196/22093

KEYWORDS

human-technology interaction; maternal health; mHealth; mobile phone; utilization; Kenya

Introduction

Background

The growth in mobile technology has led to the budding of mobile innovations such as mobile health (mHealth), whose use could solve some of the most persistent challenges in low- and middle-income countries. mHealth refers to “innovations that integrate the use of mobile and wireless devices to improve healthcare outcomes, healthcare services, and health research into care delivery” [1]. Infectious diseases and maternal health are the 2 main health outcomes where mHealth has had the greatest effect in developing countries [2]. This is no surprise because maternal health is one of the pressing needs in most resource-poor countries, with mortality rates being much higher in these countries, than in their high-income counterparts. Although maternal mortality in sub-Saharan Africa (SSA) dropped by 45% to 546 maternal deaths per 100,000 live births

XSL•FO
RenderX

between 1990 and 2015, these figures are still much higher than those for European and Commonwealth of Independent States countries that were already at a figure of 69 maternal deaths per 100,000 live births in 2009, which had dropped by more than half to 25 maternal deaths per 100,000 live births by 2015 [3].

Pregnancy is a complex period in a woman's life, and various factors influence the uptake and use of maternal health interventions to generate health outcomes. In low-resource countries, maternal health clients' perceptions of care, quality of service, sociocultural, and socioeconomic factors may all contribute to how, why, and when maternal clients use interventions. Some studies have shown that maternal clients' perceptions of health care providers may positively or negatively influence the uptake of services [4-6]. In addition, perceptions of quality of service and the level of satisfaction with quality of care may contribute to the motivation or delay in seeking care [4]. The price of services and the high direct and indirect costs of care may also impede access [7,8], although a woman's autonomy in health care decision-making may also be a factor in her financial independence. Equally, uncertainties surrounding pregnancy from sociocultural beliefs [9,10] may influence how and when maternal clients interact with maternal services. Pregnancy in most parts of SSA is a largely collectivist experience. Other family members play a role in pregnancy-related decision-making and may influence a maternal clients' use of maternal services [11,12].

Studies show that the "use of mobile technology can improve client knowledge base, service uptake and timely management of emerging pregnancy complications" [13], thereby improving maternal health outcomes. In information systems, however, it is well established that the characteristics of the technology and its context of use may influence its use in the first place [14-16]. Together with the sociocultural factors explored earlier, the interplay of factors to produce mHealth use outcomes proves complex. In a nascent field such as mHealth, whose evidence base is largely anecdotal, it is therefore useful to examine and understand the link between intervention characteristics and their use in light of contextual realities in which the interventions are implemented.

Although there have been many reviews of maternal and child health (MCH) mHealth interventions in low-and middle-income countries [17-22], most of these reviews have only studied the effectiveness of the interventions in terms of health and clinical indicators. One review [22] explored the influence of such interventions in MCH practices, such as clinic attendance and assisted delivery. Thus, to our knowledge, no review has explored the interventions' design and implementation characteristics in light of their use.

Objectives

This review seeks to contribute to exploring the influence of design and implementation characteristics on the use of MCH mHealth interventions. Unlike most reviews, we opt to adopt a country-specific analysis to allow for depth rather than breadth of analyzing mHealth interventions. Therefore, we chose a country that has a high number of implemented mHealth programs as a case, because insights from such a country may be beneficial in charting a direction for mHealth in SSA. Kenya is one of the countries whose maternal mortality is still high, ranking 19th in both SSA and the world. Owing to its concomitant growth in mobile technologies, Kenya has become a hot spot for mHealth interventions [23,24]. Mobile growth statistics show that, together with South Africa and Nigeria, Kenya's mobile economy ranks high in Africa [25]. Mobile phone ownership has grown exponentially over the past decade from 33% in 2007 to an estimated 86% in 2018 [26], with over 100% penetration rate in 2020, attributed to multiple SIM card ownership [27]. Our interest is in maternal health interventions with which the maternal health clients directly interact, rather than those delivered via a health care worker or volunteer. Hence, the high penetration of mobile phones—that provide a channel over which these mHealth interventions are delivered—made Kenya an interesting case to explore the objectives of this study.

Governments have identified the need for inventories of mHealth programs as an important prerequisite for tracking eHealth innovations in countries [28]. In many countries, the lack of clarity on what maternal mHealth interventions exist could potentially further *pilotitis* and duplication of efforts among implementers. Consequently, in addition to allowing for depth in tracing interventions, a tighter geographic focus allows for the study to contribute to developing a country-specific preliminary inventory. Although Njoroge et al [29] conducted a review in Kenya, their review was not targeted specifically at maternal health; therefore, it may not offer such an inventory of maternal mHealth interventions.

We believe that this review will complement existing studies that highlight the influence of mHealth use on MCH practices and outcomes by elucidating how the characteristics of such technologies may influence use. We think that this is important because only by their successful use will mHealth interventions achieve their lauded potential to improve maternal health in developing countries. These insights would be useful for mHealth designers and implementers and provide a direction for areas that need to be strengthened in mHealth research. The resulting inventory may also be useful to maternal mHealth implementers and the government to consider existing interventions before implementing new ones, in a bid to promote collaboration around mHealth solutions, and decrease *pilotitis*.

Methods

Overview

This study adopted a systematic review to rigorously identify and select maternal mHealth interventions to be analyzed. Many eHealth implementations in Kenya have not been reported in peer-reviewed literature [29]. Therefore, the study adopted a combination of sources to capture both peer-reviewed and gray literature. Table 1 summarizes the search strategies used. The combination of search terms from Textbox 1 to form search phrases consisted of 2 to 3 components: a word that described mHealth and related technology, a word that described maternal health and pregnancy, and the country name, that is *Kenya*, to limit the results to our geographical area of interest. We conducted the first phase of the literature search between February and April 2019 and the second phase between April

XSL•FO
RenderX

and May 2021. We have used the terms intervention and program interchangeably to describe a specific mHealth project.

Table 1. Search strategy (adapted from Njoroge et al [29]).

Step	Peer-reviewed sources	Non–peer-reviewed sources
1	Peer-reviewed sources from the databases EBSCOhost, PubMed, Scopus, Web of Science, ACM, and Google Scholar	Non–peer-reviewed sources, such as web-based portals for Kenya's most read newspapers (Nation and Standard) and organizational reports (WHO[a], mHealth[b] Alliance, and IDRC[c])
2	Manual searches of references in documents	Web portals for eHealth projects in Kenya
3	N/A[d]	Profit-based and nonprofit organizational websites
4	N/A	Personal communication with players

[a]WHO: World Health Organization.

[b]mHealth: mobile health.

[c]IDRC: International Development Research Centre.

[d]N/A: not applicable.

Textbox 1. Keywords used for the systematic literature review.

Keywords

- Kenya, mhealth and/or m-health, mobile health, maternal, maternal, neonatal and child health and/or MNCH, pregnant wom?n, pregnan*, mobile, mobile phone, mobile telephon*, innovation*, cell phone, text messag*, SMS, voice call*

Search Strategy

In our first round of searching for peer-reviewed sources, we selected literature sources, databases, websites, and registers based on their relevance and likely coverage of literature and applied the search strategy detailed in Table 1. The databases for peer-reviewed sources included EBSCOhost (capturing resources from Academic Search Premier, CINAHL, LISTA, MEDLINE, Newspaper Source, and SocINDEX), PubMed, Scopus, Web of Science, Association for Computing Machinery, and Google Scholar. Following this, we conducted a manual search using reference trailing to augment and fill in any gaps in our search strategy.

The primary author (KS) developed the search terms by reviewing previously published peer-reviewed studies. The search terms were reviewed and tested for completeness by the second author (PM). We used the same terms in both peer- and nonpeer-reviewed searches. The search terms included Boolean-paired key words, variants, and spelling variations as detailed in Textbox 1.

Our second round of search was targeted at gray sources to identify interventions that were existent but which might not have been formally evaluated. The non–peer-reviewed sources incorporated web-based portals for eHealth, profit-based and nonprofit organizational websites, newspaper articles, organization blogs, and reports. The final step, which can be deemed rather subjective, was initiated by the primary researcher through personal communication with mHealth players in Kenya, linked to the interventions retrieved from gray sources. This was done to gather missing information and validate what had been accessed from the websites, as well as to trace other programs that the researchers may have missed. To start with, the researcher contacted 2 people, who provided referrals to 2 other people, bringing the total number to 4 (Textbox 2). Interviews were conducted in person. The participants offered some high-level details of the programs, most of which had already been gathered from their websites and publicly available resources.

Textbox 2. Participants and their affiliations.

Participants and affiliations

- Participant 1 was affiliated with BabyMed.
- Participant 2 was affiliated with TotoHealth.
- Participant 3 was affiliated with Amref, Kenya.
- Participant 4 was affiliated with Amref, Kenya.

Inclusion and Exclusion Criteria

Eligible materials included journal articles, conference proceedings, and published information from governments and other organizations' portals. Peer-reviewed sources were required to have the full text available on the web for review. Gray references to interventions were included if the existence of the intervention could be established by more than 1 source or personal communication with key players or both. As reflected in government reports and documents, English is the

XSL•FO
RenderX

business language in Kenya. Having confirmed that there was a corresponding English source for the few Swahili sources that we could identify, we chose to include articles that were published in English. We did not apply any year restrictions to the search because mHealth is fairly nascent in most low-resource countries.

As a guide for the selection of maternal mHealth programs, we adopted the World Health Organization's definition of maternal health as the health of women during pregnancy, childbirth, and 6 weeks post partum. Therefore, in general, programs that addressed other areas in the reproductive, maternal, neonatal, and child health continuum were included only if they had a maternal health component delivered to maternal clients during this period. In the same manner, as HIV contributes to approximately 20% of maternal deaths [30], we included prevention of mother-to-child transmission and antiretroviral treatment adherence programs, which are initiated during pregnancy and targeted at improving pregnancy outcomes for maternal clients.

Articles were included only if the mHealth interventions they discussed were immediately and directly related to the improvement of maternal health outcomes. In addition, the peer-reviewed citations needed to have some evaluation information regarding the requisite interventions. Protocol-study dyads were included to provide a rich description of the intervention. For programs identified from gray sources, the inclusion depended on a verifiable existence, which was done by double-checking with other sources, typically by entering the intervention name as a search text on Google Search (Google, LLC) or by talking to health players (Table 1 and Textbox 2).

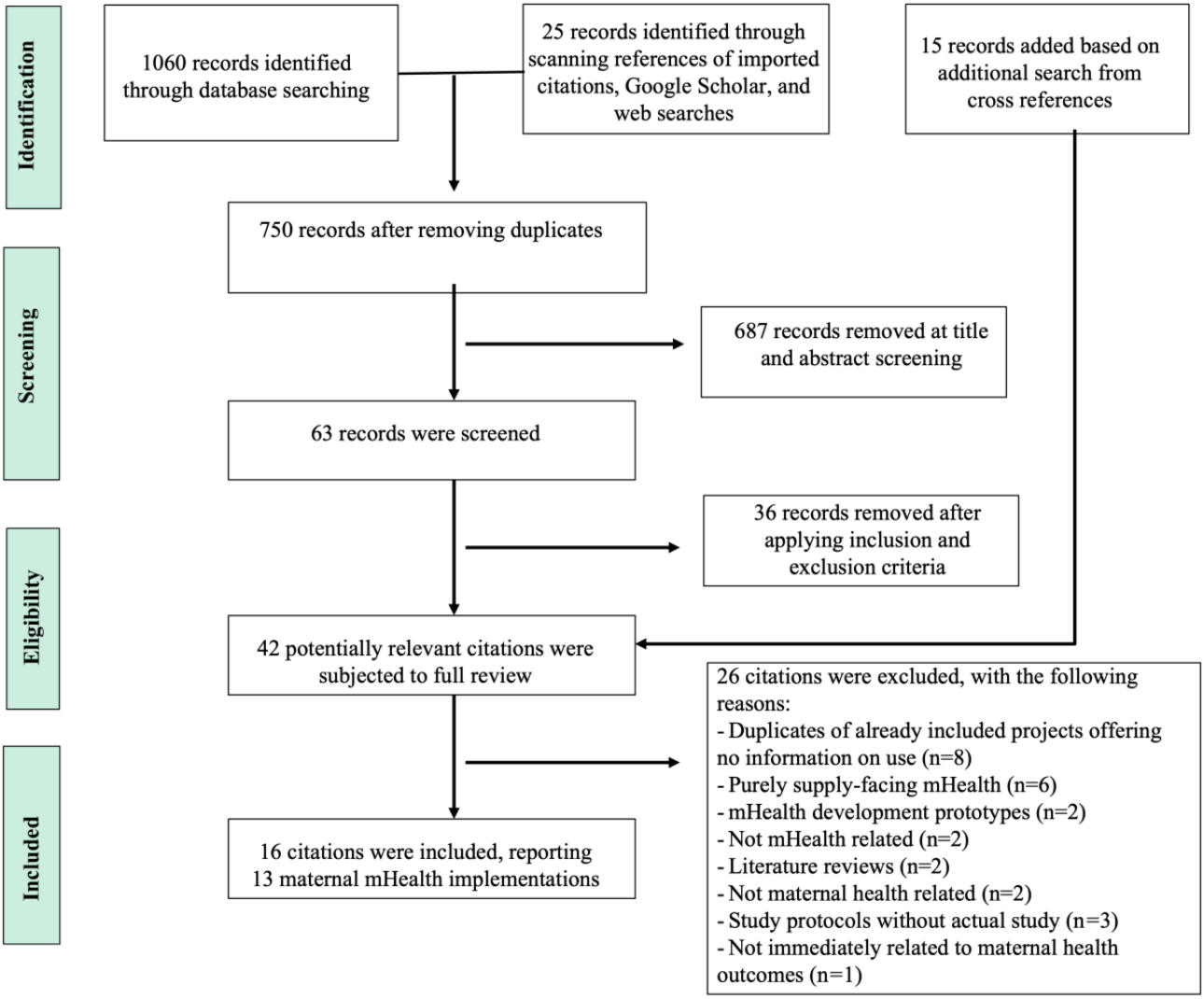
Articles on mHealth programs that did not have evidence of maternal outcomes were excluded, and so were those that lacked evidence of outcomes in Kenya. We also excluded literature reviews and studies whose main objective was to describe the development of mHealth system prototypes, without having an actual (not beta) deployment where maternal clients interacted with it. Study protocols whose evaluation outcomes could not be traced were excluded because they lacked findings on which user experiences could be assessed. Articles describing interventions that were purely supply facing were also omitted, as client experiences cannot be best explained by observing supply-side use. mHealth programs were counted only once if they were discussed in more than 1 article.

Screening Process

The primary author (KS) applied the search terms to the search sources, imported results to EndNote (Clarivate, Inc), removed duplicates, and screened for inclusion based on title and abstract, and then by skimming through the full article. To determine the final inclusion, 2 authors independently reviewed the full text of the potential citations. At this stage, citations that were supply facing, such as those targeted for use by community health workers or volunteers, were excluded. The screening process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Literature screening process. mHealth: mobile health.



Synthesis of Results

Using the search strategy detailed in Table 1, we identified 1085 citations from peer-reviewed and non–peer-reviewed sources. An additional 15 citations were retrieved through additional searches (Figure 1). After full review, a total of 16 citations were included, featuring 13 unique client-facing maternal mHealth project interventions (Tables 2 and 3).

XSL•FO
RenderX

Table 2. Aggregated characteristics of the maternal mobile health interventions in Kenya (N=13).

Implementation characteristics	Technology used, n (%)			
	SMS	SMS + voice	Other	Total
Other stakeholders[a]				
Health care providers	7 (54)	3 (23)	1 (8)	11 (85)
Others	4 (31)	0 (0)	1 (8)	5 (38)
Application area				
Education and behavior change	9 (69)	2 (15)	1 (8)	12 (92)
Others	0 (0)	1 (8)	0 (0)	1 (8)
Location				
Urban or periurban	4 (31)	2 (15)	1 (8)	7 (54)
Rural	4 (31)	1 (8)	0 (0)	5 (38)
Both	1 (8)	0 (0)	0 (0)	1 (8)

[a]Three interventions involved both health care providers and other stakeholders and were thus tallied twice.

XSL•FO
RenderX

Table 3. Design and implementation characteristics of maternal mobile health interventions in Kenya.

Author	Title	Intervention name	Design and implementation characteristics	Outcomes related to use
Perrier et al [31]	Male Partner Engagement in Family Planning SMS Conversations at Kenyan Health Clinics	Unnamed SMS platform	• Toll-free SMS • Automated weekly messages • Messages to participants with enrolled partner using inclusive wording • A question related to the message topic to encourage engagement • Home visits to reach the partners • Dedicated staff to answer the messages • Privacy in group messaging feature by sending messages separately to each person's individual phone among couple dyads	• Including the male partner engaged more households than would otherwise be included in the conversation • Use significantly dropped when intervention stopped being free • Individualized responses from study staff help build a level of trust in the SMS system opening the door to more engagement • Privacy within couple dyads encouraged conversation
Perrier et al [32]	Engaging Pregnant Women in Kenya with a Hybrid Computer-Human SMS Communication System	Mobile WACh	• Two-way SMS • Human-mediated (computer automates bulk-sending of messages and responses are tailored by a staff) • Allow unstructured messages • Send personalized time-sensitive messages • Each message salutes with mother's name • Toll-free • Dedicated nurse tells mother about intervention and enroll her, highlighting that the intervention is free • Language choices according to user • A question related to the message topic to encourage engagement.	• Unstructured messages increase access by allowing users with little experience to participate and engage • Stage personalized messages made women to feel cared for • The availability of a nurse to answer questions made women feel cared for
Ronen et al [33] and Drake et al [34][a]	SMS messaging to improve ART[b] adherence: perspectives of pregnant HIV-infected women in Kenya on HIV-related message content	Mobile WAChX	• Tailored messages based on woman's stage in the pregnancy or postpartum continuum • SMS delivered in preferred language • Message includes question related to the message topic that solicits engagement • Women in two-way arm communicate with the study nurse via SMS at any time • SMS content developed in consultation with target group • Congratulatory message sent when ANC visit is attended • Option to opt out at any point • Messages include salutation with nurse's and client's name	• Messages helped women feel cared for • Messages improved perceptions of care • Concerns about confidentiality in receiving HIV-overt content (mainly because of third-party access to their phone) • Anonymity in medium (SMS) resulted in patients feeling that they could send overt HIV messages to the nurse
Fairbanks [35]	Perceptions of SMS content for Pregnant and Postpartum Kenyan Women Infected with HIV	Mobile WACh-X	• Tailored SMS • Confidentiality (sending covert rather than explicit messages)	• Feeling cared for and supported • Improved engagement in HIV and MCH[c] health outcomes • Caring messages improve provider-patient relationships • Messages serve as a catalyst to engaging in conversation with their partners

XSL•FO
RenderX

Author	Title	Intervention name	Design and implementation characteristics	Outcomes related to use
Harrington et al [36]	An mHealth[d] SMS intervention on Postpartum Contraceptive Use Among Women and Couples in Kenya: A Randomized Controlled Trial	Mobile WACh XY	• Question at the end of message designed to promote SMS dialogue • Message content corresponding to participants' gestational age or postpartum week • Semiautomated messages with nurse's input to tailor responses to client questions • Female participants recruited by female nurses and male by male study staff • Toll-free	• Two-way SMS with a nurse, and an involved partner increased postpartum contraceptive use. • Two-way SMS results in a high level of participant engagement in SMS dialogue with study nurses
Pintye et al [37]	Two-Way Short Message Service (SMS) Communication May Increase Pre-Exposure Prophylaxis Continuation and Adherence Among Pregnant and Postpartum Women in Kenya	mWACh-PrEP	• Two-way SMS to allow real-time communication • A dedicated nurse to receive and respond to messages • SMS message development informed by theory • Toll-free SMS service • Messages include salutation with nurse's and client's name • Messages ends with a question to solicit engagement • Autonomously exit the program • Multiple language options	• Real-time communication facilitated continued pre-exposure prophylaxis use • High SMS engagement from participants in response to automated push messages • Women reported consulting by SMS with the nurse and continuing pre-exposure prophylaxis because of the nurse's advice • Diminished response to automated messages after one month
Patel et al [38]	Providing Support to Pregnant Women and New Mothers through Moderated WhatsApp Groups: a Feasibility Study	Jacaranda Health	• Moderator's participation in the service was part-time • Skilled moderator (with basic nursing background) asks questions to stimulate conversation • 10 women/group with similar gestational age • Moderator referred participants to the health facility to address individual medical questions (maintaining patient confidentiality)	• Groups created small community for women to learn from and support each other • Dissatisfaction over delayed responses from the nurse that resulted in some maternal clients abandoning the intervention
Bardosh et al [39] and Awiti et al [40][e]	Operationalizing mHealth to improve patient care: a qualitative implementation science evaluation of the WelTel texting intervention in Canada and Kenya	WelTel Kenya-2 Grand Challenges Canada	• Interactive two-way SMS with optional voice call from provider to patient • Manual messaging • Free of charge • Occasional push messages	• Two-way SMS allowed patients to seek feedback on questions and problems, giving the sense that someone cared • The two-way communication Improved relationship between patients and providers • 20% of HIV patients enrolled in the intervention immediately; 80% enrolled only after being encouraged by other patients • High number of nonrespondents that did not respond to the weekly messages
Fedha [13]	Impact of Mobile Telephone on Maternal Health Service Care: A Case of Njoro Division	Njoro Hospital	• Optional provider-patient follow-up	• No data on user experiences (evaluation done based on health outcome indicators only)
Finocchario-Kessler et al [41]	A Pilot Study to Evaluate the Impact of the HIV Infant Tracking System (HITSystem 2.0) on Priority PMTCT[f] Outcomes	HITSystem v2.0)	• Participants choose preferred message content and frequency • All message content in one language—Kiswahili	

XSL•FO
RenderX

Author	Title	Intervention name	Design and implementation characteristics	Outcomes related to use
				• Standard care participants more likely to be disengaged from care than those receiving intervention care • Higher preference for daily SMS than weekly or monthly
Jones et al [42]	A Short Message Service (SMS) increases postpartum care-seeking behavior and uptake of family planning of mothers in peri-urban public facilities in Kenya	PROMPTS	• Messages sent in Swahili to lower accessibility barriers • SMSs were free of charge • Stage-relevant messages • Messaging around postpartum checkups was broad, and explain why postpartum care was imperative	• Women who were told by the health care provider to come back for a postpartum checkup were 14 times more likely to come back than those who received messages only
Germann et al [43]	Jamii Smart| KimMNCHip–referrals, mSavings and eVouchers	KimMNCHip (Jamii Smart)	• Mobile wallet facilities allow mothers to save and manage their delivery and health care costs	• No data
BabyMed [44]	BabyMed	Baby Med	• Time-sensitive reminders sent to mothers and partners • One message per week delivered to husbands and other family members to support the mother and baby	• Received positive feedback in involving partners
Luseka et al [45]	An Evaluation of Totohealth Mobile Phone Platform on Maternal and Child Health Care in Kenya	Toto Health	• Time-sensitive and targeted information and reminders sent to mothers and their partners • Initially toll-free and later made a paid for service • Inconsistency in timing of messages • Periodic visit by a Toto health personnel to the mother	• Feeling cared for • Partners felt that they had a friend to educate them about what's happening to their wives • Abandoned used when intervention was introduced as a paid for service • Dissatisfaction with the inconsistency • Increased trust on the system whenever personnel visited the mothers

[a]The research protocol for Ronen et al [33].

[b]ART: antiretroviral therapy.

[c]MCH: maternal and child health.

[d]mHealth: mobile health.

[e]The research protocol for Bardosh et al [39].

[f]PMTCT: prevention of mother-to-child transmission.

We sought to characterize the interventions according to the technology channel in use, the involvement of other stakeholders, mHealth application area as documented in Labrique et al [46], and location of implementation, whether urban or rural (Table 2). Where an implementation could have been placed in more mHealth application areas, or was not sufficiently described to understand its content, it was placed in the category in which the researchers deemed as the best fit.

In synthesizing the articles, we focused on the design and implementation characteristics and the impact of these on the use experiences of users, as described in the citations. We also took note of the type of evaluation, for example, if it was a randomized controlled trial (RCT) or another type of evaluation.

Results

Intervention Characteristics

Table 2 describes the aggregated characteristics of the identified mHealth programs, whereas Table 3 elaborates on their design, implementation, and use details.

Use of SMS in mHealth

SMS was the predominant technology used in both urban and rural maternal mHealth implementations. Some of the most common uses of SMS include the delivery of health information and appointment reminders. Interventions implemented different SMS calibrations, including one-way (push) or two-way SMS,

XSL•FO
RenderX

most of which were computer-automated, with a human component to respond to client questions, often referred to as hybrid systems [31-33,36,37,41,42]. Fewer programs incorporated other channels such as voice [13,39], which were mostly available for the health care provider for follow-up purposes, and even fewer reported using other messaging options such as WhatsApp (Meta Platforms, Inc) [38].

Target Users and mHealth Application Areas

As expected, all interventions mainly targeted the maternal client with the aim of client education and behavior change. Most interventions also involved the health care provider to either follow up on clients or to respond to client questions [13,32,33,37,39]. In addition to involving health care workers, few interventions also involved other stakeholders in the women's life such as their partners [33,36] and other women to offer group support [38].

Implementation Location

One of the main motivations for the use of technologies such as mHealth by health programs is to extend the geographic reach of health care, particularly in resource-strained environments [47]. This may be in the form of addressing the shortage of health care providers, as well as unequal distribution of health services that may exist between social groups such as urban and rural or rich and poor. The results of this review suggest that a higher number of interventions were piloted in urban and periurban areas.

Nature of Programs and Evaluation

Most of the deployments were short-lived funded pilot studies, whereas 2 interventions [44,45] represented proprietary social enterprises that were privately owned. Almost all the interventions that had been evaluated were RCTs [32,33,36,37,42]. Of the 3 interventions whose evaluation details could not be traced, 2 (66%) were privately owned social enterprises, and one was a multi-stakeholder program [43].

Influence of Design and Implementation Characteristics on Use

Engaging Other Stakeholders May Promote Use

Many interventions involve health care providers in the implementation process to execute various roles. For example, in the Mobile WACh program [32], a nurse was assigned to tell the maternal clients about the intervention, enroll them, and highlight that the intervention was free. Most interventions that implemented two-way SMS used a dedicated health care provider to respond to maternal health client queries via SMS [31-33,37,39]. In rare cases, such as for Jacaranda Health, for implementing a moderated WhatsApp group support system, the intervention made use of part-time staff [38]. These interventions, which were integrated into mainstream care by involving health care providers at the local health facilities that the maternal client attended, resulted in improved perceptions of care and better provider-patient relationships. Fewer interventions that engaged other community members such as male partners, for example [31,45], showed that mHealth educational messages led to better health outcomes, resulting from increased engagement with the mHealth content.

Design and Implementation Characteristics May Facilitate Use

The results of our synthesis show that interventions were more readily adopted and used when they were offered free of charge [31,32,36,37,39,42]. Interventions whose content was not stage-based, such as mWACh-PrEP, experienced diminished use after some time. The diminished use was likely experienced when the users felt sufficiently onboarded regarding the logistics and continuation or discontinuation of pre-exposure prophylaxis [37]. However, when interventions delivered timely, useful, time-sensitive, stage-based information, accompanied by an appropriate *tone* and *voice*, the mothers felt *cared for* [32,33,35,39,45] and continued active use. The additional access to a health care provider to answer questions [31-33,37] resulted in the women developing positive perception toward care and toward health care providers.

The anonymity offered by SMS, as well as the anonymity in message content, especially where the target users were HIV-infected women, influenced the way maternal health clients interacted with the intervention. In Mobile WAChX [33], maternal health clients expressed concerns about confidentiality in receiving HIV-overt content, mainly because of possible third-party access to their phones. The anonymity of the SMS channel, compared with face-to-face communication, also afforded users the opportunity to engage with overt questions. In Mobile WACh, Mobile WAChX, and Mobile WACh-PrEP [32,33,37], the messages included a salutation to the maternal client using her name before the actual message content. This was also seen to improve the perception of personalized care among the maternal health clients, which further resulted in them feeling *cared for*.

Frugal Technology Such as SMS Promotes Opportunities for Use

Almost all interventions reported the use of SMS to deliver messages related to maternal health care. mHealth programs, particularly those using SMS, have been shown to increase the uptake of maternal health services in developing countries [32,36,42]. In the randomized trials, the users in the two-way SMS intervention arms showed better engagement with the mHealth intervention [31-33,36,37,39]. As these were implemented free of charge, maternal health clients were able to address health-related concerns by sending a message to the health care provider and receiving feedback in real time. Unstructured message implementations also allowed increased access and use by allowing users with little technical experience to participate and engage with the interventions [32].

Discussion

Principal Findings

The study's findings suggest that various design and implementation characteristics may influence use. From our analysis, we identified three main considerations: (1) engaging other relevant stakeholders to promote use, (2) designing interventions with characteristics that facilitate and promote use, and (3) considerations for the use of SMS technology.

XSL•FO
RenderX

Engaging the Maternal Community of Purpose in the Design Section and Implementation

Various individuals ranging from health care providers to other community members share a common interest and have stakes in women's pregnancies in most low- and middle-income countries. These stakeholders form a *community of purpose*. Involving health care providers in the implementation process may have positive outcomes regarding the use of mHealth interventions. Health care providers wield power in many health contexts, especially when they are regarded as gatekeepers of medical information. A similar observation was made in the literature, suggesting that health information technologies are likely to be more successful if providers encourage patients to use them [48]. Other findings in technology acceptance literature support that people may adopt technology based on the belief that important others think that they should do so [49,50].

Having a *human face*, such as a physician, who in the implementation context represents a trusted entity [51], could also promote adoption by minimizing the perceived risks and uncertainties about using the intervention, like clarifying the toll-free access [32]. Furthermore, having a trusted human face to introduce the intervention has been reported to avert concerns about perceived risks such as airtime loss that may prevent adoption [16], a concern that might be more pertinent in lower-income user groups. As health care providers are also considered a trustworthy source of care [51], their involvement in interventions may help legitimize mHealth information, thus averting the maternal clients' perceived risks brought about by cultural beliefs related to certain maternal care practices and habits. In support of these arguments, adoption theories have suggested that mass media alone is not enough to drive the adoption of technology [14]. Therefore, they point to the need for rich channels of communication (eg, face-to-face communication) to share information about new technology in contexts where the personal and sociocultural characteristics of the target users result in high uncertainty regarding technology. Maternal health in developing countries represents one such context, where the uncertainty in using technology may additionally be attributed to the overall uncertainties surrounding the pregnancy experience.

In addition to health care providers, various individuals share a common interest and have stakes in women's pregnancies in developing countries. Although older female relatives provide care and support, male partners are often responsible for the financial needs of the maternal clients [5,11]. During pregnancy, women especially rely on family support for responsibilities related to childcare and other areas that are considered female domains [52]. Therefore, being away from family significantly reduces family support for women [53]. The increased need for support may, therefore, promote the success of novel interventions where maternal health clients are brought together to offer group support with the direction of a trained health care provider, as seen in the study by Patel et al [38].

Engaging stakeholders such as partners may increase engagement with mHealth content because of the interdependent nature of the maternal health care–seeking context. An intervention that includes partners and other significant others in the maternal health client's life may serve to reduce the negotiation that the maternal health client must engage in to ensure her use if the intervention is a culturally appropriate behavior [51]. Rogers et al [14], in the diffusion of innovations theory, uses the term compatibility to refer to the degree to which using an innovation is perceived as consistent with the existing sociocultural values and beliefs of the adopters.

Better health outcomes may also reflect the affordance that technology offers to negotiate cultural *rules*. For example, although pregnancy is often considered a woman's domain in which men are not involved [54,55], designing interventions that involve men engenders more of their participation without causing overt disharmony in social norms. Altogether, engaging the relevant stakeholders in the design and implementation process could have positive outcomes on mHealth use because of the interdependent nature of the maternal health care–seeking context in developing countries, especially in societies that are more collectivist in nature. However, interventions also need to be aware of the complex interpersonal relationship dynamics in a maternal health context [31] when calibrating the community of purpose engagement.

Designing Interventions With Characteristics to Facilitate Use

Although the success of toll-free interventions could be linked to the socioeconomic status of maternal clients, toll-free services may also have increased the trialability of the intervention, as observed by Sowon and Chigona [16]. The trialability of an innovation is positively correlated with the likelihood of its adoption [14]. In health care, trialability is often linked to minimal financial investment [56].

The findings also suggest that the quality of mHealth information and what it evokes in users is crucial to maternal mHealth. Some researchers have suggested that when technology is *faceless*, users build trust by assessing the quality of the information [57], which is often used by mHealth users as a proxy for quality of service, especially in innovations such as mHealth, where health information is critical. Other studies have observed that mHealth may be underused when its users express low trust in their integrity and benevolence or when there is no demonstration of in-depth knowledge and clear concise information [58,59]. The additional access to health care providers provided in two-way SMS calibrations may further increase the perceptions of usefulness, thus engendering use. Subsequent responses from health care providers build positive perceptions toward the providers. Altogether, the quality of the information, perceived usefulness, and positive attitude toward providers could result in positive perceptions of care.

Other characteristics such as anonymity may positively influence use because they afford users the opportunity to engage matters of stigma or cultural taboos associated with certain conversations [60,61]. Though tailoring messages with a client's name may promote personalization, it may also thwart the potential benefits to be gained from perceived anonymity. However, because the findings show that there are other options to personalization, such as sending time-sensitive messages based on a woman's stage of pregnancy, the decision on what to personalize in an intervention should be context-dependent. Interventions need

XSL•FO
RenderX

to identify the most salient characteristics to be tailored for the delivery of health interventions [62]. As seen in the results, personalization is important because such characteristics influence the clients' perceptions of quality of service, and subsequently, their judgments on satisfaction and use. Satisfaction is necessary for the continued use of mHealth interventions [63].

Opportunities and Challenges With SMS

Our results confirm the findings in other reviews [17,21,29], indicating that most mHealth interventions in developing countries show a proliferation of interventions that uses short messaging. The popularity of SMS may be ascribed to the fact that it is accessible even on the most basic feature phones and attracts lower costs than, for example, voice calls. Most rural and underserved populations are likely be in ownership of a feature phone and not a smartphone [26]. In addition, SMS is easy to use because it does not require high literacy levels. Although the reasons for SMS popularity may be sensible, other features that allow group interaction, as demonstrated by Patel et al [38], may provide more novel solutions, especially in situations where maternal clients are already separated from their usual family support, such as in urban areas. Pure SMS interventions may also exclude those who cannot read and write, thus creating further gaps in health.

Although push messages may be less complicated for mHealth providers and designers to offer [42], they lack the robustness and flexibility that two-way SMS offers for users. One-way SMSs that allow users to submit text to a server lack a feedback loop that leaves a user wondering if their message was received [64]. On the contrary, two-way SMS interventions allow consumers to engage with care and engender better use. Some of the programs reviewed in this study reported innovative ways of engaging users in SMS. For example, posing reflective questions with most messages to solicit engagement [31-33], was seen to have positive outcomes on how maternal clients engaged with the intervention.

These findings suggest that unstructured message formats increase usability. However, such programs require human intervention because the automation of responses would be complex. Using humans to respond to client questions may create further bottlenecks, which may create dissatisfaction and limit use, as seen in the study by Patel et al [38]. In their study, maternal health clients abandoned the intervention because they were dissatisfied with the delayed responses. The period of waiting for care may have negative implications on health outcomes, as maternal health clients may engage with alternative sources of care [5]. These alternative sources may offer contradictory information to mainstream care, thus worsening health conditions. Hence, finding ways to engage users, especially given the asynchronous nature of SMS, will be critical to the long-term success of such interventions.

Conclusions and Recommendations

This review intended to provide insights on how mHealth design and implementation characteristics may influence use by reviewing and analyzing maternal mHealth interventions in Kenya. The 2016-2030 Kenya National eHealth Policy also identifies the need for mHealth inventories as a prerequisite to managing the licensing and audit of interventions by the Ministry of Health. Thus, the results of this study offer a potential maternal mHealth inventory.

The findings reveal that mHealth design and implementation characteristics play a critical role in how maternal health clients use mHealth interventions. Certain characteristics could promote the use of mHealth interventions but the causal relationship largely depends on the context, as users interact with technology within their local realities. The study identified that involving stakeholders, having characteristics that facilitate use, and how SMS is deployed in interventions are all factors that could influence use.

However, these insights are generated from evaluations that only marginally discuss experiences of use. This review reveals that most mHealth evaluations [13,32,33,40,65-68] are implemented as RCTs, which mostly evaluate maternal health interventions based on quantitative health outcome indicators. Thus, there is little evidence of studies explaining the mechanisms, that is, why, when, and how interventions work or do not work. This calls for researchers and implementers to conduct more research in this area, to understand how mHealth interventions generate outcomes, or how they are used in their relevant contexts. One way to do this is to theoretically elaborate on the findings of this study to explain the mechanisms by which the design and implementation factors produce varied mHealth use outcomes. Such studies guided by theory will make it more possible to generalize results beyond a specific context, which may help in understanding how and whether to scale interventions. Although RCTs will remain useful in assessing the effectiveness of mHealth, they will be insufficient if adopted as the only method [69]. Qualitative investigations, especially on use, will complement RCTs and provide better evidence for mHealth.

Limitations

This study had some limitations. Some of the data derived from the articles included in this study were from evaluations of mHealth interventions that did not purposefully report user experiences of the interventions. Consequently, the data may be insufficient for generalization. In addition, it proved difficult to reach the players to interview them. We believe that there may be other small-scale interventions that could have been implemented, which could have been identified only by the stakeholders involved in their implementation. As we depended largely on publicly available resources, the list of maternal mHealth implementations presented here may not be complete, and the findings are also limited to what could be accessed.

XSL•FO
RenderX

Authors' Contributions

All authors contributed to this study. KS designed the study, conducted the literature search, screened the sources for inclusion, and wrote the paper. KS and PM reviewed and synthesized the eligible citations. WC critiqued and provided guidance through the different phases of the study. All authors reviewed the final draft of the manuscript.

Conflicts of Interest

None declared.

References

1. White A, Thomas DSK, Ezeanochie N, Bull S. Health Worker mHealth Utilization: A Systematic Review. Comput Inform Nurs 2016 Mar 7. [doi: 10.1097/CIN.0000000000000231] [Medline: 26955009]
2. Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N. Barriers to the use of mobile health in improving health outcomes in developing countries: systematic review. J Med Internet Res 2019 Oct 09;21(10):e13263 [FREE Full text] [doi: 10.2196/13263] [Medline: 31593543]
3. Maternal mortality. UNICEF DATA. 2017. URL: https://data.unicef.org/topic/maternal-health/maternal-mortality/ [accessed 2019-05-22]
4. Mannava P, Durrant K, Fisher J, Chersich M, Luchters S. Attitudes and behaviours of maternal health care providers in interactions with clients: a systematic review. Global Health 2015 Aug 15;11(1):36 [FREE Full text] [doi: 10.1186/s12992-015-0117-9] [Medline: 26276053]
5. Atukunda E, Mugyenyi G, Obua C, Musiimenta A, Agaba E, Najjuma J, et al. Women's choice to deliver at home: understanding the psychosocial and cultural factors influencing birthing choices for unskilled home delivery among women in Southwestern Uganda. J Pregnancy 2020;2020:6596394 [FREE Full text] [doi: 10.1155/2020/6596394] [Medline: 32566299]
6. Uldbjerg CS, Schramm S, Kaducu FO, Ovuga E, Sodemann M. Perceived barriers to utilization of antenatal care services in northern Uganda: a qualitative study. Sex Reprod Healthc 2020 Mar;23:100464. [doi: 10.1016/j.srhc.2019.100464] [Medline: 31710878]
7. Ensor T, Cooper S. Overcoming barriers to health service access: influencing the demand side. Health Policy Plan 2004 Mar 01;19(2):69-79. [doi: 10.1093/heapol/czh009] [Medline: 14982885]
8. Doctor H, Radovich E, Benova L. Time trends in facility-based and private-sector childbirth care: analysis of Demographic and Health Surveys from 25 sub-Saharan African countries from 2000 to 2016. J Glob Health 2019 Dec;9(2):020406 [FREE Full text] [doi: 10.7189/jogh.09.020406] [Medline: 31360446]
9. Simkhada B, Teijlingen EV, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. J Adv Nurs 2008 Feb;61(3):244-260. [doi: 10.1111/j.1365-2648.2007.04532.x] [Medline: 18197860]
10. Nyemba-Mudenda M, Chigona W. mHealth outcomes for pregnant mothers in Malawi: a capability perspective. Inform Technol Develop 2017 Nov 15;24(2):245-278. [doi: 10.1080/02681102.2017.1397594]
11. Kaiser JL, Fong RM, Hamer DH, Biemba G, Ngoma T, Tusing B, et al. How a woman's interpersonal relationships can delay care-seeking and access during the maternity period in rural Zambia: an intersection of the Social Ecological Model with the Three Delays Framework. Soc Sci Med 2019 Jan;220:312-321 [FREE Full text] [doi: 10.1016/j.socscimed.2018.11.011] [Medline: 30500609]
12. Serizawa A, Ito K, Algaddal A, Eltaybe R. Cultural perceptions and health behaviors related to safe motherhood among village women in Eastern Sudan: ethnographic study. Int J Nurs Stud 2014 Apr;51(4):572-581. [doi: 10.1016/j.ijnurstu.2013.08.007] [Medline: 24054068]
13. Fedha T. Impact of mobile telephone on maternal health service care: a case of Njoro division. Open J Prevent Med 2014;04(05):365-376. [doi: 10.4236/ojpm.2014.45044]
14. Rogers E. Diffusion of Innovations. 4th ed. New York: Aimon and Schuster; 2010:1-4516.
15. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
16. Sowon K, Chigona W. Trust in mHealth: how do maternal health clients accept and use mHealth interventions? In: Proceedings of the Conference of the South African Institute of Computer Scientists and Information Technologists 2020. 2020 Presented at: SAICSIT '20: Conference of the South African Institute of Computer Scientists and Information Technologists 2020; September 14 - 16, 2020; Cape Town South Africa p. 189-197. [doi: 10.1145/3410886.3410895]
17. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
18. Nurmatov UB, Lee SH, Nwaru BI, Mukherjee M, Grant L, Pagliari C. The effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: protocol for a systematic review and meta-analysis. J Glob Health 2014 Jun;4(1):010407 [FREE Full text] [doi: 10.7189/jogh.04.010407] [Medline: 24976966]

XSL•FO
RenderX

19. Poorman E, Gazmararian J, Parker RM, Yang B, Elon L. Use of text messaging for maternal and infant health: a systematic review of the literature. Matern Child Health J 2015 May;19(5):969-989. [doi: 10.1007/s10995-014-1595-8] [Medline: 25081242]
20. Sondaal SF, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016 May;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
21. Watterson JL, Walsh J, Madeka I. Using mHealth to improve usage of antenatal care, postnatal care, and immunization: a systematic review of the literature. Biomed Res Int 2015;2015:153402 [FREE Full text] [doi: 10.1155/2015/153402] [Medline: 26380263]
22. Obasola OI, Mabawonku I, Lagunju I. A review of e-Health interventions for maternal and child health in Sub-Sahara Africa. Matern Child Health J 2015 Aug;19(8):1813-1824. [doi: 10.1007/s10995-015-1695-0] [Medline: 25652059]
23. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
24. Linde DS, Korsholm M, Katanga J, Rasch V, Lundh A, Andersen MS. One-way SMS and healthcare outcomes in Africa: systematic review of randomised trials with meta-analysis. PLoS One 2019 Jun 6;14(6):e0217485 [FREE Full text] [doi: 10.1371/journal.pone.0217485] [Medline: 31170176]
25. The mobile economy Sub-Saharan Africa 2018. GSMA. 2018. URL: https://www.marketlinks.org/resources/report-mobile-economy-sub-saharan-africa-2018 [accessed 2022-01-11]
26. Smartphone ownership is growing rapidly around the world, but not always equally internet. Pew Research Center. 2019. URL: https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center_Global-Technology-Use-2018_2019-02-05.pdf [accessed 2022-11-01]
27. Fourth quarter sector statistics report for the financial year 2019/20. Communications Authority of Kenya. 2020. URL: https://ca.go.ke/wp-content/uploads/2020/10/Sector-Statistics-Report-Q4-2019-2020.pdf [accessed 2022-01-11]
28. Kenya National eHealth Policy 2016-2030. Ministry of Health. 2016. URL: https://repository.kippra.or.ke/handle/123456789/1786 [accessed 2022-01-11]
29. Njoroge M, Zurovac D, Ogara EA, Chuma J, Kirigia D. Assessing the feasibility of eHealth and mHealth: a systematic review and analysis of initiatives implemented in Kenya. BMC Res Notes 2017 Feb 10;10(1):90 [FREE Full text] [doi: 10.1186/s13104-017-2416-0] [Medline: 28183341]
30. Lathrop E, Jamieson DJ, Danel I. HIV and maternal mortality. Int J Gynaecol Obstet 2014 Nov 17;127(2):213-215 [FREE Full text] [doi: 10.1016/j.ijgo.2014.05.024] [Medline: 25097142]
31. Perrier T, Harrington E, Ronen K, Matemo D, Kinuthia J, John-Stewart G, et al. Male partner engagement in family planning SMS conversations at Kenyan health clinics. In: Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. 2018 Presented at: COMPASS '18: ACM SIGCAS Conference on Computing and Sustainable Societies; June 20 - 22, 2018; Menlo Park and San Jose CA USA p. 1-11. [doi: 10.1145/3209811.3209857]
32. Perrier T, Dell N, DeRenzi B, Anderson R, Kinuthia J, Unger J, et al. Engaging pregnant women in Kenya with a hybrid computer-human SMS communication system. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15: CHI Conference on Human Factors in Computing Systems; April 18 - 23, 2015; Seoul Republic of Korea p. 1429-1438. [doi: 10.1145/2702123.2702124]
33. Ronen K, Unger JA, Drake AL, Perrier T, Akinyi P, Osborn L, et al. SMS messaging to improve ART adherence: perspectives of pregnant HIV-infected women in Kenya on HIV-related message content. AIDS Care 2018 Apr;30(4):500-505 [FREE Full text] [doi: 10.1080/09540121.2017.1417971] [Medline: 29254362]
34. Drake AL, Unger JA, Ronen K, Matemo D, Perrier T, DeRenzi B, et al. Evaluation of mHealth strategies to optimize adherence and efficacy of Option B+ prevention of mother-to-child HIV transmission: rationale, design and methods of a 3-armed randomized controlled trial. Contemp Clin Trials 2017 Jun;57:44-50 [FREE Full text] [doi: 10.1016/j.cct.2017.03.007] [Medline: 28315480]
35. Fairbanks J. Perceptions of SMS content for pregnant and postpartum Kenyan women infected with HIV. Dissertation, University of Washington. 2017. URL: https://digital.lib.washington.edu/researchworks/handle/1773/40183 [accessed 2021-04-20]
36. Harrington EK, Drake AL, Matemo D, Ronen K, Osoti AO, John-Stewart G, et al. An mHealth SMS intervention on postpartum contraceptive use among women and couples in Kenya: a randomized controlled trial. Am J Public Health 2019 Jun;109(6):934-941. [doi: 10.2105/AJPH.2019.305051] [Medline: 31067089]
37. Pintye J, Rogers Z, Kinuthia J, Mugwanya KK, Abuna F, Lagat H, et al. Two-way short message service (SMS) communication may increase pre-exposure prophylaxis continuation and adherence among pregnant and postpartum women in Kenya. Glob Health Sci Pract 2020 Mar 30;8(1):55-67 [FREE Full text] [doi: 10.9745/GHSP-D-19-00347] [Medline: 32139420]
38. Patel SJ, Subbiah S, Jones R, Muigai F, Rothschild CW, Omwodo L, et al. Providing support to pregnant women and new mothers through moderated WhatsApp groups: a feasibility study. Mhealth 2018;4:14 [FREE Full text] [doi: 10.21037/mhealth.2018.04.05] [Medline: 29963559]

XSL•FO
RenderX

39. Bardosh K, Murray M, Khaemba A, Smillie K, Lester R. Operationalizing mHealth to improve patient care: a qualitative implementation science evaluation of the WelTel texting intervention in Canada and Kenya. Global Health 2017 Dec 06;13(1):87 [FREE Full text] [doi: 10.1186/s12992-017-0311-z] [Medline: 29208026]
40. Awiti PO, Grotta A, van der Kop M, Dusabe J, Thorson A, Mwangi J, et al. The effect of an interactive weekly mobile phone messaging on retention in prevention of mother to child transmission (PMTCT) of HIV program: study protocol for a randomized controlled trial (WELTEL PMTCT). BMC Med Inform Decis Mak 2016 Jul 11;16(1):86 [FREE Full text] [doi: 10.1186/s12911-016-0321-4] [Medline: 27401475]
41. Finocchario-Kessler S, Brown M, Maloba M, Nazir N, Wexler C, Goggin K, et al. A pilot study to evaluate the impact of the HIV Infant Tracking System (HITSystem 2.0) on Priority Prevention of Mother-to-Child Transmission (PMTCT) outcomes. AIDS Behav 2021 Aug 11;25(8):2419-2429. [doi: 10.1007/s10461-021-03204-0] [Medline: 33709212]
42. Jones RM, Kimenju G, Subbiah S, Styles A, Pearson N, Rajasekharan S. A short message service (SMS) increases postpartum care-seeking behavior and uptake of family planning of mothers in peri-urban public facilities in Kenya. PLoS One 2020 Sep 30;15(9):e0239213 [FREE Full text] [doi: 10.1371/journal.pone.0239213] [Medline: 32997684]
43. Germann S. Jamii Smart| KimMNCHip - referrals, mSavings and eVouchers. Digital Square. 2020. URL: https://lib.digitalsquare.io/handle/123456789/77642 [accessed 2022-01-11]
44. BabyMed. URL: https://www.babymed.com/ [accessed 2022-01-11]
45. Luseka B. An evaluation of Toto-health mobile phone platform on maternal and child health care in Kenya. Dissertation, University of Nairobi. 2018. URL: http://erepository.uonbi.ac.ke/handle/11295/105904 [accessed 2021-04-14]
46. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
47. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. Bull World Health Organ 2012 May 1;90(5):332-340 [FREE Full text] [doi: 10.2471/BLT.11.099820] [Medline: 22589566]
48. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. J Med Internet Res 2013 Feb;15(2):e43 [FREE Full text] [doi: 10.2196/jmir.2243] [Medline: 23470453]
49. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
50. Li X, Hess TJ, Valacich JS. Why do we trust new technology? A study of initial trust formation with organizational information systems. J Strat Inform Syst 2008 Mar;17(1):39-71. [doi: 10.1016/j.jsis.2008.01.001]
51. Cheruto SK. How the healthcare-seeking socio-cultural context shapes maternal health clients? mHealth utilisation in a Kenyan context. Dissertation, University of Cape Town. 2021. URL: https://open.uct.ac.za/handle/11427/34016 [accessed 2022-01-11]
52. Clark S, Madhavan S, Cotton C, Beguy D, Kabiru C. Who helps single mothers in Nairobi? The role of kin support. J Marriage Fam 2017 Aug 28;79(4):1186-1204 [FREE Full text] [doi: 10.1111/jomf.12404] [Medline: 29479116]
53. Madhavan S, Clark S, Araos M, Beguy D. Distance or location? How the geographic distribution of kin networks shapes support given to single mothers in urban Kenya. Geograph J 2017 Oct 06;184(1):75-88. [doi: 10.1111/geoj.12230]
54. Mumtaz Z, Salway S. Gender, pregnancy and the uptake of antenatal care services in Pakistan. Sociol Health Illn 2007 Jan;29(1):1-26 [FREE Full text] [doi: 10.1111/j.1467-9566.2007.00519.x] [Medline: 17286703]
55. Aarnio P, Kulmala T, Olsson P. Husband's role in handling pregnancy complications in Mangochi District, Malawi: a call for increased focus on community level male involvement. Sex Reprod Healthc 2018 Jun;16:61-66. [doi: 10.1016/j.srhc.2018.02.005] [Medline: 29804778]
56. Cain M, Mittman R. Diffusion of innovation in health care. In: iHealth Reports. Oakland, CA: California Healthcare Foundation; 2002.
57. Beldad A, de Jong M, Steehouder M. How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. Comput Hum Behav 2010 Sep;26(5):857-869. [doi: 10.1016/j.chb.2010.03.013]
58. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. Info Systems J 2018 Jan 29;28(6):995-1019. [doi: 10.1111/isj.12179]
59. Sillence E, Briggs P, Harris P, Fishwick L. Health Websites that people can trust – the case of hypertension. Interact Comput 2007 Jan;19(1):32-42. [doi: 10.1016/j.intcom.2006.07.009]
60. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep) 2008 Nov(175):1-1422. [Medline: 19408968]
61. Dean AL, Makin JD, Kydd AS, Biriotti M, Forsyth BWC. A pilot study using interactive SMS support groups to prevent mother-to-child HIV transmission in South Africa. J Telemed Telecare 2012 Oct;18(7):399-403. [doi: 10.1258/jtt.2012.120118] [Medline: 23034933]
62. Ryan P, Lauver DR. The efficacy of tailored interventions. J Nurs Scholarsh 2002 Dec;34(4):331-337. [doi: 10.1111/j.1547-5069.2002.00331.x] [Medline: 12501736]

XSL•FO
RenderX

63. Akter S, D'Ambra J, Ray P, Hani U. Modelling the impact of mHealth service quality on satisfaction, continuance and quality of life. Behav Inform Technol 2013 Dec;32(12):1225-1241. [doi: 10.1080/0144929x.2012.745606]
64. Hellström J, Karefelt A. Participation through mobile phones: a study of SMS use during the Ugandan general elections 2011. In: Proceedings of the Fifth International Conference on Information and Communication Technologies and Development. 2012 Presented at: ICTD '12: Fifth International Conference on Information and Communication Technologies and Development; March 12 - 15, 2012; Atlanta Georgia USA p. 249-258. [doi: 10.1145/2160673.2160705]
65. Harrington EK. Evaluation of an mHealth SMS dialogue strategy to meet women's and couples' postpartum contraceptive needs in Kenya (Mobile WACh XY): a randomized controlled trial. Doctoral Dissertation. 2018. URL: https://digital.lib.washington.edu/researchworks/handle/1773/41691 [accessed 2022-01-11]
66. McConnell M, Ettenger A, Rothschild CW, Muigai F, Cohen J. Can a community health worker administered postnatal checklist increase health-seeking behaviors and knowledge?: evidence from a randomized trial with a private maternity facility in Kiambu County, Kenya. BMC Pregnancy Childbirth 2016 Jun 04;16(1):136 [FREE Full text] [doi: 10.1186/s12884-016-0914-z] [Medline: 27260500]
67. Mushamiri I, Luo C, Iiams-Hauser C, Ben Amor Y. Evaluation of the impact of a mobile health system on adherence to antenatal and postnatal care and prevention of mother-to-child transmission of HIV programs in Kenya. BMC Public Health 2015 Feb 07;15:102 [FREE Full text] [doi: 10.1186/s12889-015-1358-5] [Medline: 25886279]
68. Odeny TA, Onono M, Owuor K, Helova A, Wanga I, Bukusi EA, et al. Maximizing adherence and retention for women living with HIV and their infants in Kenya (MOTIVATE! study): study protocol for a randomized controlled trial. Trials 2018 Jan 29;19(1):77 [FREE Full text] [doi: 10.1186/s13063-018-2464-3] [Medline: 29378622]
69. Hatcher AM, Bonell CP. High time to unpack the 'how' and 'why' of adherence interventions. AIDS 2016 May 15;30(8):1301-1304. [doi: 10.1097/QAD.0000000000001071] [Medline: 27128329]

Abbreviations

MCH: maternal and child health
mHealth: mobile health
RCT: randomized controlled trial
SSA: sub-Saharan Africa

Edited by G Eysenbach; submitted 03.07.20; peer-reviewed by B Thies, H Pratomo; comments to author 25.12.20; revised version received 16.09.21; accepted 27.11.21; published 27.01.22.

Please cite as:
Sowon K, Maliwichi P, Chigona W
The Influence of Design and Implementation Characteristics on the Use of Maternal Mobile Health Interventions in Kenya: Systematic Literature Review
JMIR Mhealth Uhealth 2022;10(1):e22093
URL: https://mhealth.jmir.org/2022/1/e22093
doi:10.2196/22093
PMID:35084356

©Karen Sowon, Priscilla Maliwichi, Wallace Chigona. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 27.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Impact of Smartphone App–Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials

Maria J Serrano-Ripoll[1,2,3*], PhD; Rocío Zamanillo-Campos[1,2*], PhD; Maria A Fiol-DeRoque[1,2], PhD; Adoración Castro[1,4], PhD; Ignacio Ricci-Cabello[1,2,5], PhD

[1]Health Research Institute of the Balearic Islands (IdISBa), Palma de Mallorca, Spain
[2]Balearic Islands Health Services, Primary Care Research Unit of Mallorca, Palma de Mallorca, Spain
[3]Primary Care Prevention and Health Promotion Network, RedIAPP, Barcelona, Spain
[4]Research Institute of Health Sciences (IUNICS), University of the Balearic Islands, Palma de Mallorca, Spain
[5]Centro de Investigación Biomedica en Red (CIBERESP), Madrid, Spain
[*]these authors contributed equally

Corresponding Author:
Maria J Serrano-Ripoll, PhD
Balearic Islands Health Services, Primary Care Research Unit of Mallorca
Escola Graduada, 3
Palma de Mallorca, 07002
Spain
Phone: 34 971175883
Email: mjserranor@yahoo.es

Abstract

Background: Depression is a serious, disabling mental disorder that severely affects quality of life. Patients with depression often do not receive adequate treatment. App-based psychotherapy is considered to have great potential to treat depression owing to its reach and easy accessibility.

Objective: We aim to analyze the impact of app-based psychological interventions for reducing depressive symptoms in people with depression.

Methods: We conducted a systematic literature review and meta-analysis. We searched Medline, Embase, PsycINFO, Web of Science, and Cochrane Central Register of Controlled Trials from inception to December 23, 2020. We selected randomized controlled trials to examine the impact of app-based psychological interventions for reducing depressive symptoms in people with depression. Study selection, data extraction, and critical appraisal (using the Cochrane Risk of Bias tool for randomized studies and the ROBINS-I tool for nonrandomized studies) were conducted independently by 2 reviewers. Where possible, we pooled data using random effects meta-analyses to obtain estimates of the effect size of the intervention. We conducted post hoc meta-regression analyses to explore the factors associated with intervention success.

Results: After screening 3468 unique references retrieved from bibliographic searches and assessing the eligibility of 79 full texts, we identified 12 trials (2859 participants) evaluating 14 different interventions. Of 14 trials, 7 (58%) were conducted in the United States; 3 (25%) trials, in Asia (Japan, South Korea, and China); 1 (8%) trial, in Australia; and 1 (8%) trial, in Germany. Of the 12 trials, 5 (42%) trials presented a low risk of bias. The mean duration of the interventions was 6.6 (SD 2.8) weeks. Two-thirds of the interventions were based on cognitive behavioral therapy alone or included it in combination with cognitive control therapy, positive psychology, brief behavioral activation, or mindfulness- and acceptance-based therapy. With no evidence of publication bias, a pooled analysis of 83% (10/12) of the trials and 86% (12/14) of the interventions showed that app-based interventions, compared with a control group receiving usual care or minimal intervention, produced a moderate reduction in depressive symptoms (standardized mean difference [SMD] −0.51, 95% CI −0.69 to −0.33; 2018/2859, 70.58% of the participants; I^2=70%). Our meta-regression analyses indicated that there was a greater reduction in symptoms of depression (P=.04) in trials that included participants with moderate to severe depression (SMD −0.67, 95% CI −0.79 to −0.55), compared with trials with participants exhibiting mild to moderate depression (SMD −0.15, 95% CI −0.43 to −0.12).

XSL•FO
RenderX

Conclusions: App-based interventions targeted at people with depression produce moderate reductions in the symptoms of depression. More methodologically robust trials are needed to confirm our findings, determine which intervention features are associated with greater improvements, and identify those populations most likely to benefit from this type of intervention.

Trial Registration: PROSPERO CRD42019145689; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=145689

(JMIR Mhealth Uhealth 2022;10(1):e29621) doi:10.2196/29621

KEYWORDS

smartphone technology; mental health interventions; depression; eHealth; mHealth; apps; systematic review; meta-analysis; mobile phone

Introduction

Background

Worldwide, approximately 350 million people are affected by depression [1]. In 2010, it was estimated to be the second largest contributor to the global disease burden [2], and by 2030, it is expected to become the leading contributor [3]. Depression is a highly prevalent condition that affects approximately 4.4% of the world's population [4]. It can have a negative impact on one's mood and cause emotional distress, potentially interfering with daily functioning [5]. Symptoms of depression range in severity (mild to severe) and duration (months to years) [6]. Depression is the leading contributor to suicide, accounting for approximately 800,000 deaths per year [7]. There is an increasing number of people living with depression worldwide, especially in low-income countries [8,9]. Even in high-income countries, most patients with depression do not receive treatment [10].

The digital market is full of apps designed to improve the mental health of people with depression; however, most of them remain untested in clinical trials and suffer from numerous limitations, such as being designed without content based on the recommendations of experts [11,12]. Therefore, there is a potential risk in the use of such apps, as their therapeutic benefits have not been proven. A recent review of apps targeting depression and anxiety-related conditions [13] observed that the techniques used by some apps were not based on evidence, and in some cases, the manifestation of the techniques promoted by apps could be potentially harmful.

Despite the proliferation of systematic reviews examining the impact of mobile health (mHealth) interventions on mental health during the last decade, the available base of evidence concerning the impact of mobile apps for treating people with depression is still weak. Most of the available reviews offer an overview of the impact of mental mHealth interventions but do not focus on their specific impact on depression [14-18]. A small portion of reviews specifically examine the impact of mHealth interventions on depression; however, some of them rely on user evaluations rather than on evidence from trials [13,17,19]. Although a recent review [20] examined the impact of mHealth interventions on depression, most of the included interventions targeted other mental health problems (insomnia, bipolar disorder, anxiety, and amnesia, among others). To the best of our knowledge, no previous systematic review of randomized controlled trials (RCTs) has evaluated the impact of mHealth interventions specifically designed to improve depressive symptoms in people with depression.

Notwithstanding the above, the use of mHealth technologies for the treatment of symptoms of depression remains very attractive, as such technologies could offer potential benefits in terms of patient autonomy, the prevention of relapse, and lowering costs [14,20]. Community health representatives perceive mHealth technologies as adequate tools for actively involving patients in the management of chronic diseases [21]. Apart from intrinsic barriers to treatment, such as availability, affordability, and time constraints, people's attitudes also play an important role in non–treatment-seeking behavior [22]. Several barriers that limit the acceptability and adherence to traditional, face-to-face psychotherapy have been described, including the low self-perceived need for treatment, low mental health literacy, high self-reliance, and fear of stigmatization [22,23]. App-based psychological interventions are attractive because of their potential to overcome these barriers.

Objectives

The aim of this systematic review is to analyze the impact of app-based psychological interventions designed to reduce depressive symptoms in people with depression.

Methods

Overview

We conducted a systematic literature review following the Cochrane recommendations [24]. We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for planning, conducting, and reporting this study [25]. The review protocol was registered with PROSPERO (CRD42019145689).

Data Sources and Searches

We designed specific search strategies for biomedical and behavioral science databases (MEDLINE, Embase, PsycInfo, CINAHL, Web of Science, and Cochrane Central Register of Controlled Trials) and combined Medical Subject Headings terms and free-text keywords (Multimedia Appendix 1). We searched the databases from inception to December 23, 2020. We used EndNote X8 to create a bibliographical database and Rayyan to screen relevant records [26].

Inclusion and Exclusion Criteria

We included empirical studies examining the impact of app-based psychological interventions delivered through smartphones and aimed at reducing depressive symptoms in

XSL•FO
RenderX

people with depression compared with a nonactive control group (ie, treatment as usual, waiting-list control, or where minimal intervention was used to ensure blinding or masking). In terms of participants, we included studies involving participants with depressive symptoms of all ages and education levels as assessed using a structured clinical interview conducted according to internationally recognized standards (eg, the International Statistical Classification of Diseases and Related Health Problems and the Diagnostic and Statistical Manual of Mental Disorders) or the presence of significant depressive symptoms established using a validated screening measure (eg, the Patient Health Questionnaire and the Beck Depression Inventory). In terms of the intervention, we included studies that evaluated psychological interventions delivered through an app aimed at reducing depressive symptoms. Although multifaceted interventions were considered, to be included the app needed to have been the main component of the interventions, which were included regardless of the therapeutic orientation upon which they were based. In terms of outcomes, we included studies examining the impact of the intervention on depression severity, as measured using structured clinical interviews or validated screening measures. We included RCTs that were individually randomized and cluster randomized. We included studies in English and Spanish. Letters were excluded from the editor, editorials, study protocols, and conference abstracts. We excluded studies with intervention periods <2 weeks (as we consider this to be the minimum time necessary for changes in depressive symptoms to occur) and those with <50 randomized participants (to minimize the risk of bias arising from small sample sizes [27]).

Study Selection

In all, 2 of the 4 reviewers (MJSR, MAFD, RZC, and AC) screened all titles and abstracts for potentially eligible papers and subsequently assessed full-text papers against the eligibility criteria. They were blinded to each other's decisions. All disagreements were resolved by reaching a consensus or by involving a third reviewer.

Data Extraction and Quality Assessment

In all, 2 of the 4 reviewers (MJSR, MAFD, RZC, and AC) independently extracted quantitative data with respect to the outcomes and characteristics of the studies and interventions in the included papers. Information was extracted and entered into a standardized Microsoft Excel spreadsheet. Discrepancies among the data extractors were discussed until a consensus was reached. We contacted the authors of the included papers to request additional data when needed.

We extracted information concerning the characteristics of the trials (study design, sample size, country, setting, participants, and type of comparator), intervention (length, frequency of use, and psychological theories or techniques used), and outcomes (changes in overall depression). In all, 2 of the 4 reviewers (MJSR, MAFD, RZC, and AC) independently assessed the risk of bias in the studies selected for the meta-analyses using the Cochrane Risk of Bias tool [28]. Discrepancies were discussed among peers to reach a consensus.

Data Synthesis and Analysis

We conducted a narrative and tabulated synthesis of the findings of the included studies. We pooled data to summarize the progress made in depressive symptoms throughout the intervention and compared interventions to their relevant comparator groups. We anticipated that the included trials would vary in their settings, methods, and designs. Therefore, we used a random effects model to pool the data. Patient-reported measures for depression vary from trial to trial; therefore, we used the Cohen method to calculate pooled effect sizes based on standardized mean differences (SMDs). When needed, we reversed the scale scores (by multiplying them by −1), so that higher scores consistently conveyed higher levels of depression at all scales.

When the SD of the change between baseline and postintervention levels was not reported for either the intervention or the control group, we derived them from baseline and final SDs, assuming a degree of correlation of 0.5. Heterogeneity was quantified using the I^2 statistic, and I^2>50% was considered evidence of substantial heterogeneity. The sources of heterogeneity were explored using the Galbraith plots. Publication bias was examined using funnel plots, and the presence of asymmetry was assessed using the Begg [29] and Egger [30] tests. Meta-analyses were conducted with STATA (version 12.0; StataCorp), using the command *metan*. We conducted a range of exploratory post hoc subgroup and bivariate meta-regression analyses to explore the factors that may affect the effectiveness of smartphone interventions. On the basis of the available evidence, we decided to analyze the following potential moderators: participants' depression severity (mild to moderate vs moderate to severe) [31,32], therapeutic approaches (cognitive behavioral therapy [CBT] vs CBT plus other approaches vs behavioral activation) [20], intervention duration (1-7 vs 8-12 weeks) [31], comparator (usual care vs minimal intervention) [15,32], components of the intervention (multifaceted vs single-component interventions) [31], communication directionality (unidirectional vs bidirectional communication) [20], and the method used to assess depression (diagnostic instrument vs validated self-reported measure).

Transparency

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as it was planned have been explained (and, if relevant, reported).

Results

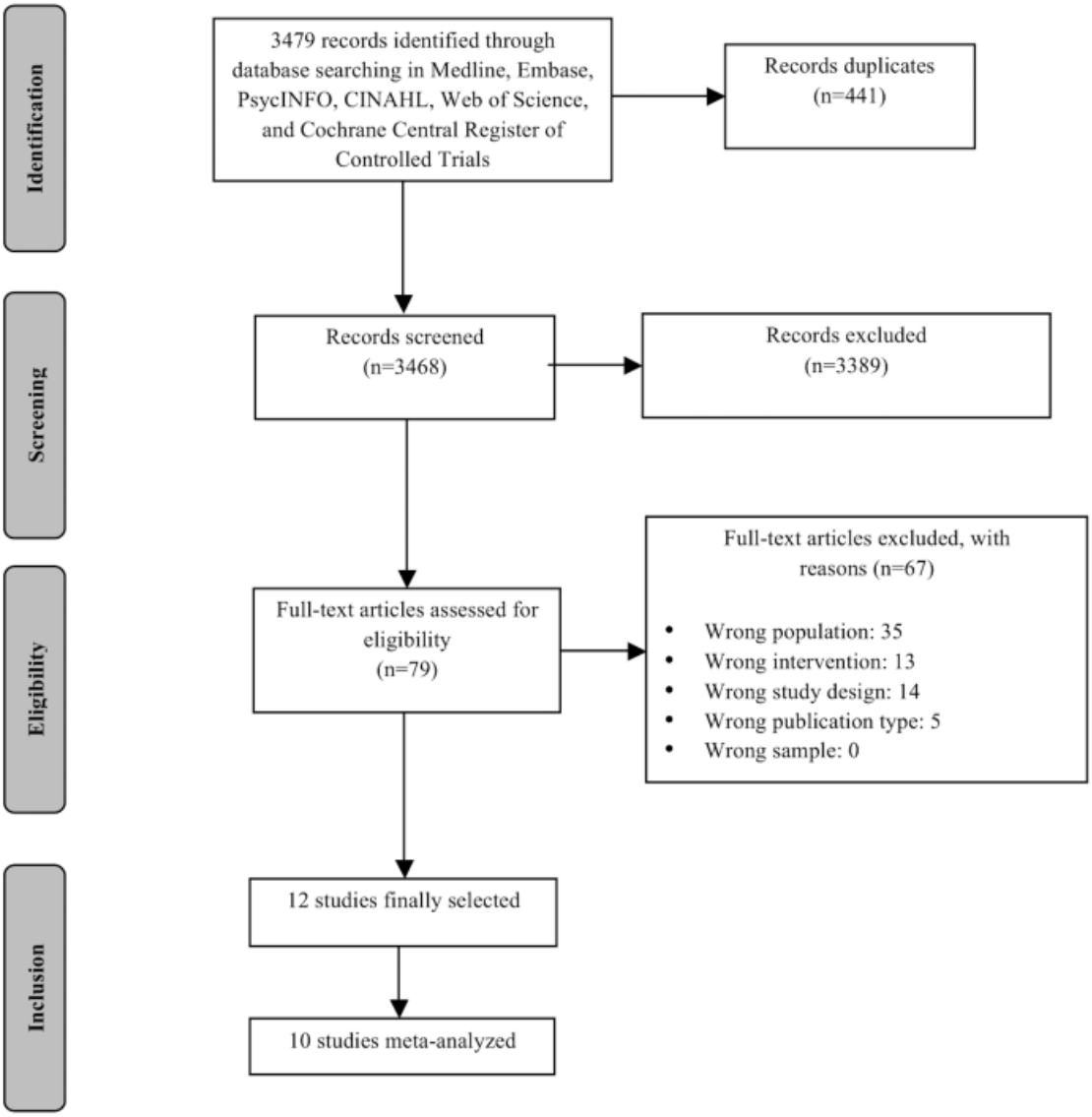
Search Results

Our search results are summarized in the following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1). Our initial search identified a total of 3468 unique citations. Screening the titles and abstracts of these studies resulted in the inclusion of 79 citations for further review. After full-text reviews, 12 trials evaluating 14 different interventions were included in the present systematic review [33-44].

XSL•FO
RenderX

Of these, 83% (10/12) of the trials were included in the meta-analysis and 17% (2/12) of the trials were excluded from the meta-analysis owing to a lack of available data.

Figure 1. PRISMA (Preferred Reporting Item for Systematic Reviews and Meta-Analyses) flow diagram.



Characteristics of the Included Studies and Interventions

A detailed description of the characteristics of the included trials is provided in Table 1 and in Multimedia Appendix 2 [33-44]. All trials consisted of individually RCTs. In all, 25 % (3/12) of the trials included people with mild to moderate depression [34,39,41], 17% (2/12) of the trials included people with mild to severe depression [33,40], 42% (5/12) of the trials included people with moderate to severe depression [35,36,42-44], 8% (1/12) of the trials included people with a diagnosis of major depression [38], and 8% (1/12) of the trials included people with a self-reported need for help with their depressive symptoms [37]. The total combined sample size for all included trials was 2859 participants. The mean (SD) number of participants per trial was 238 (182), ranging from 52 to 626. In all, 58% (7/12) of the trials were conducted in the United States [33-35,39-41,43], with 8% (1/12) each in Japan [38], Korea [36], Australia [42], Germany [37], and China [44]. A total of 42% (5/12) of the trials took place in a community setting [20,33,39,40,42], 42% (5/12) in hospitals or health organizations [34,36-38,44], and 17% (2/12) in a primary care setting [35,43]. The primary outcome of all included trials was a reduction in depressive symptoms (Table 2). A total of 58% (7/12) of the studies [34,37-39,42-44] compared the intervention against a waiting-list control group, whereas the remaining 42% (5/12) of the trials [33,35,36,40,41] compared the app intervention to another app or a waiting-list control group.

XSL•FO
RenderX

Table 1. Characteristics of included trials.

Characteristics	Value
Year the study was published (N=12), n (%)	
2015-2016	4 (33)
2017-2018	3 (25)
2019-2020	5 (42)
Number of participants, N; mean (SD; range)	2859; 238 (182; 52-626)
Age of participants (years), mean (SD)	36.12 (10.21)
Gender of participants (N=2859), n (%)	
Male	912 (31.8)
Female	1899 (66.4)
Others	48 (1.67)
Country (N=12), n (%)	
United States	7 (58.3)
Japan	1 (8.3)
South Korea	1 (8.3)
Australia	1 (8.3)
Germany	1 (8.3)
China	1 (8.3)
Instrument used to measure depression[a] (N=12), n (%)	
Patient Health Questionnaire-9	6 (50)
Patient Health Questionnaire-8	3 (25)
Beck Depression Inventory-II	2 (16.7)
Depression, Anxiety, Stress Scale-21	1 (8.3)
Center for Epidemiological Studies-Depression Scale	2 (16.7)
Setting (N=12), n (%)	
Community	5 (41.7)
Hospital or health organizations	6 (50)
Primary care	1 (8.3)
Type of approach or psychotherapy[a] (N=12), n (%)	
Cognitive behavioral therapy	8 (66.7)
Cognitive control therapy	2 (16.7)
Brief behavioral activation	1 (8.3)
Positive psychology	2 (16.7)
Mindfulness	1 (8.3)
Acceptance-based therapy	1 (8.3)
Duration of intervention in weeks, mean (SD; range)	6.6 (2.8; 4-12)

[a]Percentages exceeding 100% as categories are not mutually exclusive.

XSL•FO
RenderX

Table 2. Summary of findings.

Study	Severity of depression and instrument (cut point)	Intervention A (n) and intervention B (n)	Comparator (n)	Length	Study design	Main results
Arean et al [33]	• Mild to severe depression without suicidal ideation • PHQ-9[a] (score >5 or score on item 10 ≥2)	• EVO app (N=221) • iPST[b] app (N=209)	• Usual care (N=206)	4 weeks	Effectiveness	• No significant differences observed between the 2 interventions compared with the control after the intervention and at follow-up. • Moderately depressed participants had a greater response to Project: EVO (28/56, 50%) and iPST (39/79, 49%) than the control arm (24/76, 32%; $\chi 2$=6.46; *P*=.04) in remission rates.
Birney et al [34]	• Mild to moderate depression • PHQ-9 (score of 10-19)	• MoodHacker (N=150)	• Minimal intervention (N=150)	6 weeks	Efficacy	• Compared with the control group, the MoodHacker app had significant effects on symptoms of depression in users (*P*=.01; partial $\eta 2$=0.021) after the intervention period.
Dahne et al [35]	• Moderate to severe depression without suicidal ideation • PHQ-8[c] (score >10) and BDI-II[d] (score >13)	• Moodivate app (N=24) • Moodkit app (N=19)	• Minimal intervention (N=9)	8 weeks	Efficacy	• Over time and compared with the control group, participants using either app provided evidence of significant decreases in depressive symptoms that were sustained over the trial period.
Ham et al [36]	• Moderate to severe depression • BDI-II (score >16)	• HARUToday (N=21)	• HARUCard (attention control group) (N=21) • Waiting list (N=21)	10 weeks	Effectiveness	• BDI-II scores of the HARUToday group decreased significantly after the intervention compared with the attention control (HARUCard) and waiting-list control groups (*P*=.01).
Lüdtke et al [37]	• Subjective need for help with depressive symptoms • PHQ-9 (N/A[e])	• Be Good to Yourself (N=45)	• Usual care (waiting list) (N=45)	4 weeks	Efficacy	• Depressive symptoms decreased in both groups after the intervention period, without significant differences among groups (*P*=.95).
Mantani et al [38]	• Diagnosed major depression • PRIME-MD[f] and BDI-II (score ≥10)	• Kokoro-app (N=81)	• Usual care (N=83)	9 weeks	Effectiveness	• The intervention group improved significantly compared with the control group (95% CI 1.23-3.72; *P*<.001; SMD[g] 0.40). The benefits were maintained during the follow-up period.
Moberg et al [39]	• Mild to moderate depression • PHQ-8 (score between 5 and 14)	• Pacifica app (N=253)	• Usual care (waiting list) (N=247)	4 weeks	Effectiveness	• Participants in the intervention group demonstrated significantly greater decreases in depression. The Group x Time interaction effect size is as follows: Cohen d 0.54; *P*<.001. Rates of clinical significance change after the intervention: Pacifica, 42% (33/79); waiting list 17% (17/101); *P*<.001.
Pratap et al [40]	• Clinically significant depressive symptoms • PHQ-9 (score ≥5 or score on item 10 ≥2)	• EVO (N=83) • iPST (N=112)	• Minimal intervention (daily health tips; N=79)	4 weeks	Efficacy	• No significant differences were observed in depression outcomes among the 3 groups.

XSL•FO
RenderX

Study	Severity of depression and instrument (cut point)	Intervention A (n) and intervention B (n)	Comparator (n)	Length	Study design	Main results
Roepke et al [41]	• Mild to moderate depression • CES-D[h] (score ≥16)	• CBT-PPT[i] SB[j] (N=93) • General SB (N=97)	• Usual care (waiting list) (N=93)	4 weeks	Effectiveness	• After treatment and during follow-up, General SB participants saw greater reductions in depression scores than the control group (*P*<.001).
Tighe et al [42]	• Moderate to severe depression • PHQ-9 (score >10)	• Ibobbly (N=31)	• Usual care (waiting list) (N=30)	6 weeks	Effectiveness	• The app group showed statistically significant reductions in depression scores compared with the control group (*P*=.007).
Graham et al [43]	• Moderate to severe depression • PHQ-8 (score >10)	• IntelliCare platform (N=74)	• Usual care (waiting list) (N=72)	8 weeks	Effectiveness	• IntelliCare participants achieved greater reductions in depression and higher odds of recovery compared with the controls (odds ratio 3.25; 95% CI 1.54-6.86).
Guo et al [44]	• Moderate to severe depression • CES-D (score ≥16)	• Run Love (Wechat platform; N=150)	• Usual care (waiting list) (N=150)	12 weeks	Effectiveness	• The intervention group saw significantly reduced depression severity compared with the control group (from 23.9 to 17.7 vs from 24.3 to 23.8; mean difference −5.77, 95% CI −7.82 to −3.71; *P*<.001).

[a]PHQ-9: Patient Health Questionnaire-9.
[b]iPST: Problem-solving therapy app.
[c]PHQ-8: Patient Health Questionnaire-9.
[d]BDI-II: Beck Depression Inventory.
[e]N/A: not applicable.
[f]PRIME-MD: Primary Care Evaluation of Mental Disorders.
[g]SMD: standardized mean difference.
[h]CES-D: Center for Epidemiologic Studies-Depression Scale.
[i]CBT-PPT: cognitive behavioral therapy and positive psychotherapy.
[j]SB: SuperBetter.

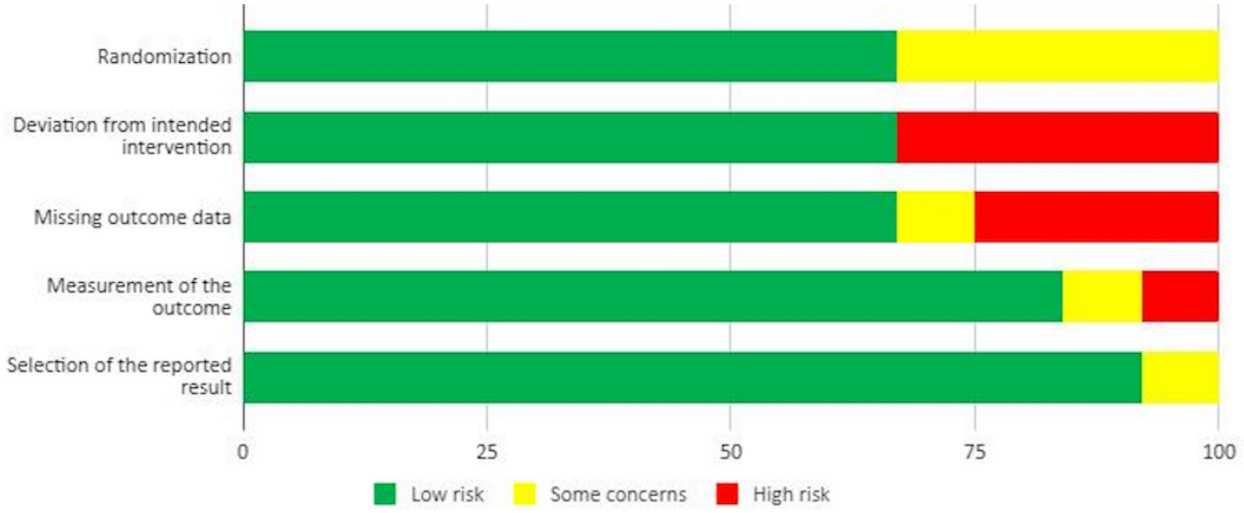
The mean duration of the interventions was 6.6 (SD 2.8) weeks, with a range of 4-12 weeks. Two-thirds of the interventions were based on CBT alone or CBT in combination with cognitive control therapy, positive psychology, brief behavioral activation, or mindfulness and acceptance-based therapy. There was variability in terms of use recommendations, with participants being recommended daily or almost daily use in 67% (8/12) of the trials [33-37,39-41] and receiving no use recommendations in 33% (4/12) of trials [38,42-44].

Risk of Bias

The results of the general risk of bias assessment are shown in Figure 2. In all, 42% (5/12) of the included studies showed a low risk of bias, 17% (2/12) of the studies showed a low risk of bias in 4 of the 5 domains considered, and the remaining 42% (5/12) of the studies showed an unclear risk. The most frequent biases included the following domains: deviation from intended intervention (high risk in 4/12, 33% studies), randomization (some concerns in 4/12, 33% studies), missing outcome data (high risk in 3/12, 25% studies and some concerns in 1/12, 8% study), and measurement of the outcome (high risk in 1/12, 8% study and some concerns in 1/12, 8% study). Our assessment of the risk of bias in individual studies is shown in Multimedia Appendix 3 [33-44].

XSL•FO
RenderX

Figure 2. Risk of bias assessment.



Impact of the Apps

In terms of impact, in the only trial that included patients with clinically diagnosed major depression [38], the authors reported that a CBT-based app intervention (Kokoro-App) improved depressive symptoms in users when compared with a waiting-list control group. A similar beneficial effect was observed in most (10/14, 71%) of the remaining app interventions: a CBT-based app (MoodHacker) [34] improved symptoms of depression compared with accessing relevant internet sites about depression. An app based on CBT and mindfulness (Pacifica App) significantly reduced depressive symptoms in users when compared with a waiting-list control group [39]. An app called SuperBetter (SB) based on CBT and positive psychotherapy strategies SB and an app that focused on self-esteem and acceptance (General SB) produced greater reductions in depression scores in users than in waiting-list participants [41]. In all, 14% (2/14) of the apps based on brief behavioral activation and CBT (Moodivate and Moodkit, respectively) produced significant decreases in depressive symptoms when compared with usual care [35], whereas another similar app (HARUToday app) [36] was also shown to have significantly reduced depressive symptoms compared with both a minimal intervention and a waiting-list control group. In a remote community setting [42], an app based on acceptance-based therapy (Ibobbly app) significantly reduced symptoms of depression in users compared with waiting-list participants. A platform containing 5 clinically focused CBT– and positive psychology–based apps [43] produced larger reductions in symptoms of depression and higher recovery rates than those seen in waiting-list participants. An app based on cognitive behavioral stress management and automatic progress monitoring (Run4Love app) [44] significantly reduced depression severity in users compared with a waiting-list control group.

Few of the interventions, however, did not consistently demonstrate the intended effect: a trial comparing 2 active interventions (Project EVO, a cognitive control app and iPST, a problem - solving therapy app) against a minimal intervention control group [33] observed that both apps had a greater effect on mood in users than the control group. However, when the same 2 interventions were subsequently evaluated in a separate trial with a high proportion of Hispanic and Latino participants [40], no significant differences were observed. A trial comparing the effect of the app Be Good to Yourself (based on CBT and mindfulness) in users with a waiting-list control group found that depressive symptoms decreased in both groups, with no significant between-group differences [37].

App Use

In all 50% (6/12) of the trials reported results concerning the app use levels. Across these studies, the data were reported using a number of different metrics (eg, percentage of participants who completed the intervention activities, number of downloads, and average use time), hindering our attempts to pool it.

App use varied widely across studies: 17% (2/12) of the trials reported that around 80% of their participants used the app as instructed (in a study by Tighe et al [42], 34/40, 85% of the participants completed all the activities and in a study by Graham et al [43], 119/146, 81.5% of the participants had some app use). However, app use was significantly lower in 3 trials: Arean et al [33] reported that 57.9% (243/420) of participants did not download the app, Dahne et al [35] reported that 43% (9/21) of participants used the app the number of times required, and Roepke et al [41] reported that 15% of the participants downloaded the app or used it to the complete content (Multimedia Appendix 3).

App use was associated with higher levels of depression at the baseline [33]. A dose-response effect was examined in 17% (2/12) of the studies: in Moberg et al [39], no significant association between overall app engagement (defined as the total number of log-ins) and symptom improvement was observed, whereas in Roepke et al [41], participants who actually downloaded General SB or the complete CBT and positive psychotherapy content achieved a significantly greater decrease in depressive symptoms.

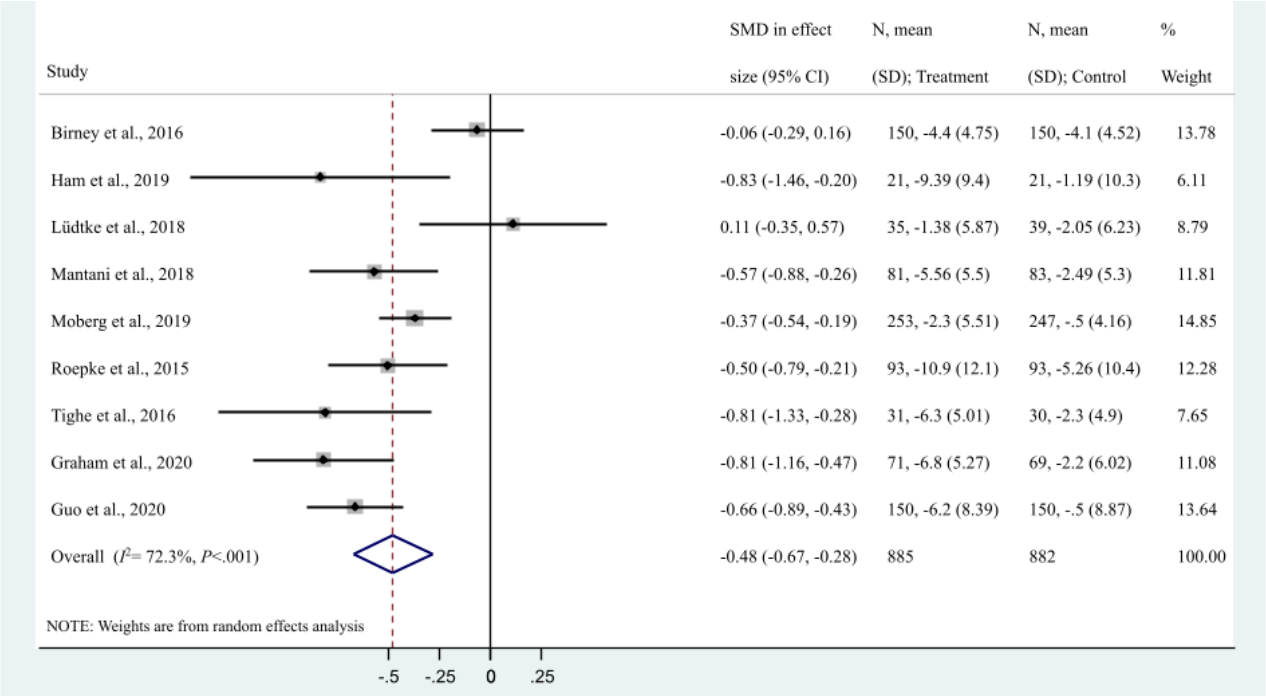
The Pooled Effects of Smartphone Interventions for Reducing Depressive Symptoms

We pooled data from 10 trials that assessed 12 interventions (Figure 3) [34-39,41-44]. Data from the remaining 17% (2/12)

XSL•FO
RenderX

of the trials included in our review were not available despite our attempts to contact the authors. According to a random effects meta-analysis, the interventions had a statistically significant and moderate effect in reducing depressive symptoms compared with control conditions in which participants received usual care or minimal intervention (SMD −0.51, 95% CI −0.69 to −0.33; 2018/2859, 70.58% of the participants; $P<.001$; $I^2=70\%$). In a sensitivity analysis that excluded from the meta-analysis, the 2 trials that most contributed to the high levels of observed heterogeneity (ie, Dahne et al [35] and Lüdtke et al [37]), the pooled impact of the interventions was greater (SMD −0.61, 95% CI −0.74 to −0.48; 1644/2859, 57.5% of the participants; $P<.001$; $I^2=34\%$). Begg and Egger tests suggested the absence of publication bias in both meta-analyses ($P=.53$ and $P=.89$, respectively, for the main meta-analysis; and $P=.31$ and $P=.93$, respectively). In a second sensitivity analysis excluding 33% (4/12) of the trials with high risk of bias, a moderate statistically significant effect was still observed (SMD −0.41, 95% CI −0.71 to −0.10; 781/2859, 27.31% of the participants; $P=.009$; $I^2=71.6\%$), with the absence of publication bias according to Egger test ($P=.53$) and Begg test ($P=.88$).

Figure 3. Effect of apps on depressive symptoms compared with active treatment and control conditions. CBT-PPT: cognitive behavioral therapy and positive psychotherapy; SB: SuperBetter; SMD: standardized mean difference.



According to our post hoc subgroup analyses (Table 3), the interventions led to better results in trials focusing on moderate to severe depression symptomatology (6/12, 50% of the trials and 9/14, 64% of the interventions [36,38,41-44]; SMD −0.67; 95% CI −0.79 to −0.55; 1144/2859, 40.01% of the participants; $I^2=0.0\%$) compared with trials involving patients with mild to moderate symptoms of depression (3/12, 25% of the trials and 4/14, 28% of the interventions [34,37,39]; SMD −0.15, 95% CI −0.43 to 0.12; 874/2859, 30.59% of the participants; $I^2=69.3\%$). This subgroup difference was statistically significant according to our meta-regression analysis ($P=.003$). The effects of interventions versus usual care (SMD −0.58, 95% CI −0.76 to −0.40) were greater than the effects of interventions versus an active control group receiving minimal intervention (SMD 0.11, 95% CI: −0.32 to 0.10). However, this difference was not statistically significant according to meta-regression ($P=.076$). The differences among the remaining subgroups were smaller and not statistically significant.

XSL•FO
RenderX

Table 3. Subgroup post hoc analyses.

Characteristics	Interventions (n=14), n (%)	Sample size (smartphone/control)	Meta-analysis, SDM[a] (95% CI)	Heterogeneity		Between-group tests, coefficient (95% CI)
				I^2 (%)	*P* value	
Depression severity						−0.49 (−0.76 to −0.22)
Mild to moderate	3 (21)	438/436	−0.16 (−0.43 to 0.12)	69.3	.04	
Moderate to severe	9 (64)	587/557	−0.67 (−0.79 to −0.55)	0.0	.68	
Type of psychotherapy						0.08 (−0.14 to 0.30)
CBT[b]	4 (28)	271/263	−0.63 (−0.81 to −0.46)	0.0	.85	
CBT + positive psychology	4 (28)	411/405	−0.55 (−0.95 to −0.15)	87.1	<.001	
CBT + mindfulness	2 (14)	288/286	−0.18 (−0.64 to 0.28)	72.7	.06	
Behavioral activation	2 (14)	55/39	−0.65 (−1.09 to −0.21)	2.2	.31	
Intervention duration						−0.24 (−0.64 to 0.16)
1-7 weeks	6 (43)	659/652	−0.41 (−0.68 to −0.14)	80.5	<.001	
8-12 weeks	6 (43)	366/341	−0.66 (−0.81 To −0.50)	0.0	.80	
Comparator						−0.41 (−0.87 to 0.05)
Usual care	9 (64)	832/825	−0.58 (−0.76 to −0.40)	62.9	.001	
Active control (minimal intervention)	3 (21)	193/168	−0.11 (−0.32 to 0.10)	0.0	.56	
Intervention components						−0.07 (−0.52 to 0.37)
Unifaceted	7 (50)	552/520	−0.48 (−0.72 to −0.24)	63.2	.01	
Multifaceted	5 (35)	473/473	−0.56 (−0.87 to −0.24)	80.3	<.001	
Directionality						−0.29 (−0.70 to 0.11)
Unidirectional communication	8 (57)	702/670	−0.41 (−0.64 to −0.17)	72.8	<.001	
Bidirectional communication	4 (28)	323/323	−0.68 (−0.84 to −0.52)	0.0	.72	
Method for assessing depression						0.07 (−0.68 to 0.81)
Diagnostic instrument	1 (7)	81/83	−0.57 (−0.88 to −0.26)	N/A[c]	N/A	
Validated self-reported measure	11 (78)	944/911	−0.50 (−0.70 to −0.30)	72.4	<.001	

[a]SMD: standardized mean difference.

[b]CBT: cognitive behavioral therapy.

[c]N/A: not applicable (subgroup with only 1 study).

Discussion

Principal Findings

In this systematic review and meta-analysis, we identified 12 RCTs examining the impact of 14 smartphone apps specifically designed to reduce depressive symptoms in people with depression. We observed that 71% (10/14) of the interventions led to a significant reduction in depressive symptoms. Our pooled analyses suggest that they had a moderate effect, which was significantly larger in interventions targeted to patients with more severe depression.

Comparison of the Main Findings With Previous Reviews

All the studies identified in our review have been published within the last 5 years, which underscores the increasing interest in this type of intervention. However, despite a growing number of studies, the available evidence base is limited by the methodological quality of the trials that have been conducted to date, most of which suffer from moderate or substantial risk of bias. This finding is in line with the results of another recent review that concluded that there is still not enough evidence to support the prescription of independent mHealth tools for depression as an adjunctive treatment [13]. Indeed, the difference between the high volume of commercially available apps and the low number of tested, evidence-based apps is striking.

The significant effects observed in our systematic review generally support the findings of previous, broader reviews [14,16,20]. However, in our review, which for the first time, meta-analyzed interventions specifically designed to reduce depressive symptoms, the observed effect size (0.51) was larger than in previous meta-analyses (ranging from 0.33 to 0.38) [14,16,20]. This may be explained by the fact that, contrary to

XSL•FO
RenderX

previous meta-analyses, we only included trials comparing the intervention with a control group that received usual care or minimal intervention. Our findings also support the results of a recent systematic review of smartphone apps for depression, which included both observational and experimental studies [45] and observed a decline in depressive symptoms in all the included studies. They additionally collected information on the attitudes of health care professionals, observing that, although they are open to therapeutic app use, professionals have limited knowledge and experience in this field.

Regarding the target population, we observed larger effects in interventions targeting people with moderate to severe depression, whereas in the review by Firth et al [20], the authors observed that mobile apps only reduced depressive symptoms in people with mild to moderate symptoms, with no differences observed in people with major depression. This difference between our review and the review by Firth et al [20] may be partially explained by the larger number of interventions we identified that were targeted to people with severe symptoms of depression (9 trials vs the 2 trials included in the review by Firth et al [20]). In their review of interventions for a broad range of mental conditions, Weisel et al [16] found that app interventions had a significant effect compared with controls for general depression but only when the comparator was a control group receiving usual care.

A recent clinical trial conducted by our team found that a psychoeducational intervention delivered through an app produced significant improvements in the mental health of health care workers on the frontline of the COVID-19 pandemic who were receiving psychotherapy or taking psychotropic drugs [46,47]. Mobile apps present numerous unique advantages, including increased accessibility to the intervention (ubiquitous access) and the fact that they provide access to people who do not seek help for their mental health problems. Thus, mHealth interventions could address the main barriers to help-seeking behaviors, such as geographic location and the stigma associated with mental illness [22,48]. Apps also provide opportunities for users to access the intervention several times a day and when it is most needed [12]. Considering their potential to improve access to mental health services and as many people do not feel the need for treatment [22], apps may be able to motivate users to seek a diagnosis or treatment, as evidenced by an app for the evaluation of depression. In this sense, participants in 92% (11/12) of the studies in the present review were encouraged to use the mobile app several times a week to daily, in some cases stimulating use with reminders. However, the available data suggest that app use is generally low (around 80% of the participants used the app as instructed in 2/12, 17% of the studies, whereas in 4/12, 33% of the studies use was <50%) concerning app use suggested from the data provided, it can be inferred that there have been few downloads of the apps, that those who downloaded them, the use has been limited and that a greater number of apps does not translate into significant improvements in depressive symptoms.

It seems that the emerging use of apps to take care of people's mental health is unstoppable, whether it is partially or in combination with the intervention of a therapist [16,17]. The evidence from the present review and meta-analysis suggests that interventions delivered via smartphones have a beneficial effect on depressive symptoms. Understanding which psychological interventions delivered through smartphones are the best and what types of patients they can best serve will require more research. Embedding process evaluations in future RCTs would provide information on mechanisms of action and a better understanding of the contexts and premises under which mHealth interventions produce beneficial effects.

New technologies are increasingly present in our lives, and mental health is not an exception. As more mental health apps are created, we will need to focus on tailoring them to more personalized populations and users so that they are likely to be more effective. Future studies should explore reliable frameworks for making use of mental health apps in the context of psychological and psychiatric care.

Strengths and Limitations of the Review

To the best of our knowledge, this is the first systematic review and meta-analysis to examine the impact of mHealth interventions specifically designed for people with depression. The strengths of this review are the large number of bibliographic databases searched, the fact that study eligibility, data extraction, and risk of bias assessments were conducted by independent senior reviewers, and the statistical analyses adhered to best-practice recommendations [24]. The current systematic review is not without limitations. Our bibliographic searches were restricted to publications in English and Spanish. In addition, we did not search for unpublished data. Both aspects may have hindered our ability to identify additional relevant trials. The differences in severity of depression, the time of the treatment received, and the differences among the studies made it difficult to establish the most effective individual components (active ingredients) of the included interventions. These differences are also likely to have contributed to the substantial heterogeneity observed in the meta-analysis (I^2=70%). However, the heterogeneity was reduced to 34% in a sensitivity analysis, excluding the 17% (2/12) of the trials that most contributed to this high level of heterogeneity. The results of the sensitivity analysis support the finding that these interventions have a moderate effect. The use of medication in addition to psychological treatment may also influence treatment outcomes, but we were not able to explore this in the review. Finally, we acknowledge the following two deviations from our published protocol: (1) the inclusion of studies that assessed depression using self-reported tools rather than diagnostic instruments (as we only identified 1 trial using a diagnostic instrument), and (2) the exclusion of studies with intervention periods <2 weeks or with <50 participants.

Conclusions

mHealth interventions targeted at people with symptoms of depression produce moderate reductions in these symptoms, with larger effects being seen in people with more severe symptoms. Although the available evidence seems to follow this line, there is still insufficient evidence to support the prescription of mHealth tools to improve depressive symptoms or as an adjunct treatment. Future research should focus on conducting more clinical trials with solid methodological

XSL•FO
RenderX

foundations to investigate the impact of digital psychological interventions for the treatment of depression.

Acknowledgments

The review has been funded by the Research Commission on Primary Care of Majorca (Spain). MJSR, MAFD, and AC have a grant from the Folium Program for Postdoctoral researchers (FOLIUM17/10, FOLIUM 19/05, and FOLIUM 19/03). Sources of funding had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Authors' Contributions

MJSR and IRC contributed to the overall study design. IRC designed and executed the bibliographic searches. MJSR, RZC, MAFD, and AC screened and selected the articles and conducted data extraction and quality appraisal. IRC conducted the statistical analyses. MJSR and IRC drafted the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 19 KB - mhealth_v10i1e29621_app1.docx]

Multimedia Appendix 2
Characteristics of the study interventions.
[DOCX File , 27 KB - mhealth_v10i1e29621_app2.docx]

Multimedia Appendix 3
Quality assessment of included studies.
[DOCX File , 17 KB - mhealth_v10i1e29621_app3.docx]

References

1. Depression: fact sheets. World Health Organization. 2021. URL: https://www.who.int/en/news-room/fact-sheets/detail/depression [accessed 2022-01-05]
2. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med 2013 Nov 5;10(11):e1001547 [FREE Full text] [doi: 10.1371/journal.pmed.1001547] [Medline: 24223526]
3. World Health Assembly. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level: report by the Secretariat. World Health Organization. 2012. URL: https://apps.who.int/iris/handle/10665/78898 [accessed 2022-01-05]
4. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [accessed 2022-01-05]
5. Moreno-Agostino D, Wu Y, Daskalopoulou C, Hasan MT, Huisman M, Prina M. Global trends in the prevalence and incidence of depression:a systematic review and meta-analysis. J Affect Disord 2021 Feb 15;281:235-243. [doi: 10.1016/j.jad.2020.12.035] [Medline: 33338841]
6. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Washington, DC: American Psychiatric Association Publishing; 2013.
7. World Health Organization. Public Health Action for the Prevention of Suicide: A Framework. Geneva: World Health Organization; 2012.
8. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006 Nov 28;3(11):e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
9. Richardson RA, Keyes KM, Medina JT, Calvo E. Sociodemographic inequalities in depression among older adults: cross-sectional evidence from 18 countries. Lancet Psychiatry 2020 Aug;7(8):673-681 [FREE Full text] [doi: 10.1016/S2215-0366(20)30151-6] [Medline: 32711708]
10. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry 2016 May;3(5):415-424. [doi: 10.1016/s2215-0366(16)30024-4]
11. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2(1):18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]

XSL•FO
RenderX

12. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 07;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
13. Baumel A, Torous J, Edan S, Kane JM. There is a non-evidence-based app for that: a systematic review and mixed methods analysis of depression- and anxiety-related apps that incorporate unrecognized techniques. J Affect Disord 2020 Aug 01;273:410-421. [doi: 10.1016/j.jad.2020.05.011] [Medline: 32560936]
14. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
15. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 09;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
16. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019 Dec 02;2(1):118 [FREE Full text] [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
17. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
18. Menon V, Rajan TM, Sarkar S. Psychotherapeutic applications of mobile phone-based technologies: a systematic review of current research and trends. Indian J Psychol Med 2017 Jan 01;39(1):4-11 [FREE Full text] [doi: 10.4103/0253-7176.198956] [Medline: 28250552]
19. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviewsuser experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
20. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct 21;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
21. Broderick A, Haque F. Mobile health and patient engagement in the safety net: a survey of community health centers and clinics. Issue brief (Commonw Fund) 2015;9:1-9. [doi: 10.15868/socialsector.25031]
22. Mojtabai R, Olfson M, Sampson NA, Jin R, Druss B, Wang PS, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med 2011 Aug;41(8):1751-1761 [FREE Full text] [doi: 10.1017/S0033291710002291] [Medline: 21134315]
23. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
24. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M. Cochrane Handbook for Systematic Reviews of Interventions. Hoboken, New Jersey, United States: John Wiley & Sons; 2019.
25. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-269 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
26. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
27. Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. PLoS One 2018 Sep 13;13(9):e0204056 [FREE Full text] [doi: 10.1371/journal.pone.0204056] [Medline: 30212588]
28. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J 2011 Oct 18;343(oct18 2):d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
29. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101. [Medline: 7786990]
30. Egger M, Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Br Med J 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
31. Park C, Zhu J, Man RH, Rosenblat J, Iacobucci M, Gill H, et al. Smartphone applications for the treatment of depressive symptoms: a meta-analysis and qualitative review. Ann Clin Psychiatry 2020 Feb;32(1):48-68. [Medline: 31675391]
32. Wu A, Scult MA, Barnes ED, Betancourt JA, Falk A, Gunning FM. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. NPJ Digit Med 2021 Feb 11;4(1):20 [FREE Full text] [doi: 10.1038/s41746-021-00386-8] [Medline: 33574573]
33. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]

XSL•FO
RenderX

34. Birney AJ, Gunn R, Russell JK, Ary DV. Moodhacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial. JMIR Mhealth Uhealth 2016 Jan 26;4(1):e8 [FREE Full text] [doi: 10.2196/mhealth.4231] [Medline: 26813737]
35. Dahne J, Lejuez C, Diaz VA, Player MS, Kustanowitz J, Felton JW, et al. Pilot Randomized Trial of a Self-Help Behavioral Activation Mobile App for Utilization in Primary Care. Behav Ther 2019 Jul;50(4):817-827 [FREE Full text] [doi: 10.1016/j.beth.2018.12.003] [Medline: 31208690]
36. Ham K, Chin S, Suh YJ, Rhee M, Yu E, Lee HJ, et al. Preliminary results from a randomized controlled study for an app-based cognitive behavioral therapy program for depression and anxiety in cancer patients. Front Psychol 2019 Jul 25;10:1592 [FREE Full text] [doi: 10.3389/fpsyg.2019.01592] [Medline: 31402881]
37. Lüdtke T, Pult LK, Schröder J, Moritz S, Bücker L. A randomized controlled trial on a smartphone self-help application (Be Good to Yourself) to reduce depressive symptoms. Psychiatry Res 2018 Nov;269:753-762. [doi: 10.1016/j.psychres.2018.08.113] [Medline: 30273901]
38. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
39. Moberg C, Niles A, Beermann D. Guided self-help works: randomized waitlist controlled trial of Pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. J Med Internet Res 2019 Jun 08;21(6):e12556 [FREE Full text] [doi: 10.2196/12556] [Medline: 31199319]
40. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]
41. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
42. Tighe J, Shand F, Ridani R, Mackinnon A, De La Mata N, Christensen H. Ibobbly mobile health intervention for suicide prevention in Australian Indigenous youth: a pilot randomised controlled trial. BMJ Open 2017 Jan 27;7(1):e013518 [FREE Full text] [doi: 10.1136/bmjopen-2016-013518] [Medline: 28132007]
43. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. JAMA Psychiatry 2020 Sep 01;77(9):906-914 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.1011] [Medline: 32432695]
44. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-Based Intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. J Med Internet Res 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: 10.2196/16715] [Medline: 32044751]
45. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. Eur Arch Psychiatry Clin Neurosci 2020 Mar 3;270(2):139-152. [doi: 10.1007/s00406-018-0974-3] [Medline: 30607530]
46. Serrano-Ripoll MJ, Ricci-Cabello I, Jiménez R, Zamanillo-Campos R, Yañez-Juan AM, Bennasar-Veny M, et al. Effect of a mobile-based intervention on mental health in frontline healthcare workers against COVID-19: protocol for a randomized controlled trial. J Adv Nurs 2021 Jun 06;77(6):2898-2907 [FREE Full text] [doi: 10.1111/jan.14813] [Medline: 33675247]
47. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A mobile phone-based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021 May 18;9(5):e27039 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]
48. Araya R, Zitko P, Markkula N, Rai D, Jones K. Determinants of access to health care for depression in 49 countries: a multilevel analysis. J Affect Disord 2018 Jul;234:80-88. [doi: 10.1016/j.jad.2018.02.092] [Medline: 29524750]

Abbreviations

CBT: cognitive behavioral therapy
mHealth: mobile health
RCT: randomized controlled trial
SB: SuperBetter
SMD: standardized mean difference

XSL•FO
RenderX

Edited by L Buis; submitted 03.05.21; peer-reviewed by M Domhardt, B Renn, R Kok, J Linardon; comments to author 06.07.21; revised version received 16.09.21; accepted 11.11.21; published 27.01.22.

Please cite as:
Serrano-Ripoll MJ, Zamanillo-Campos R, Fiol-DeRoque MA, Castro A, Ricci-Cabello I
Impact of Smartphone App–Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials
JMIR Mhealth Uhealth 2022;10(1):e29621
URL: https://mhealth.jmir.org/2022/1/e29621
doi:10.2196/29621
PMID:35084346

©Maria J Serrano-Ripoll, Rocío Zamanillo-Campos, Maria A Fiol-DeRoque, Adoración Castro, Ignacio Ricci-Cabello. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 27.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Knowledge and Expectations of Hearing Aid Apps Among Smartphone Users and Hearing Professionals: Cross-sectional Survey

Jae Sang Han[1], MD; Yong-Ho Park[2,3], MD, PhD; Jae-Jun Song[4], MD, PhD; Il Joon Moon[5], MD, PhD; Woojoo Lee[6], PhD; Yoonjoong Kim[2], MD; Young Sang Cho[5], MD; Jae-Hyun Seo[1*], MD, PhD; Moo Kyun Park[7,8*], MD, PhD

[1]Department of Otolaryngology-Head and Neck Surgery, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

[2]Department of Otolaryngology-Head and Neck Surgery, College of Medicine, Chungnam National University, Daejeon, Republic of Korea

[3]Brain Research Institute, College of Medicine, Chungnam National University, Daejeon, Republic of Korea

[4]Department of Otorhinolaryngology-Head and Neck Surgery, Korea University College of Medicine, Seoul, Republic of Korea

[5]Department of Otorhinolaryngology-Head and Neck Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[6]Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[7]Department of Otorhinolaryngology, Head & Neck Surgery, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Republic of Korea

[8]Sensory Organ Research Institute, Seoul National University, Medical Research Center, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Jae-Hyun Seo, MD, PhD
Department of Otolaryngology-Head and Neck Surgery
College of Medicine
The Catholic University of Korea
222, Banpo-daero, Seocho-gu
Seoul, 06591
Republic of Korea
Phone: 82 2 2258 6210
Fax: 82 2 595 1354
Email: revivalseo@catholic.ac.kr

Abstract

Background: Despite the increasing prevalence of hearing loss, the cost and psychological barriers to the use of hearing aids may prevent their use in individuals with hearing loss. Patients with hearing loss can benefit from smartphone-based hearing aid apps (SHAAs), which are smartphone apps that use a mobile device as a sound amplifier.

Objective: The aim of this study is to determine how ear, nose, and throat outpatients perceive SHAAs, analyze the factors that affect their perceptions, and estimate the costs of an annual subscription to an app through a self-administered questionnaire survey of smartphone users and hearing specialists.

Methods: This study used a cross-sectional, multicenter survey of both ear, nose, and throat outpatients and hearing specialists. The questionnaire was designed to collect personal information about the respondents and their responses to 18 questions concerning SHAAs in five domains: knowledge, needs, cost, expectations, and information. Perception questions were rated on a scale of 1 (strongly disagree) to 5 (strongly agree). Questions about the expected cost of SHAAs were included in the questionnaire distributed to hearing experts.

Results: Among the 219 smartphone users and 42 hearing specialists, only 8 (3.7%) respondents recognized SHAAs, whereas 18% (47/261) of respondents reported considering the use of an assistive device to improve their hearing capacity. The average perception score was 2.81 (SD 1.22). Among the factors that shaped perceptions of SHAAs, the needs category received the lowest scores (2.02, SD 1.42), whereas the cost category received the highest scores (3.29, SD 1.14). Age was correlated with the information domain (P<.001), and an increased level of hearing impairment resulted in significantly higher points in the needs category (P<.001). Patients expected the cost of an annual app subscription to an SHAA to be approximately US $86, and the

XSL•FO
RenderX

predicted cost was associated with economic status (*P*=.02) and was higher than the prices expected by hearing specialists (*P*<.001).

Conclusions: Outpatients expected SHAAs to cost more than hearing specialists. However, the perception of the SHAA was relatively low. In this regard, enhanced awareness is required to popularize SHAAs.

(JMIR Mhealth Uhealth 2022;10(1):e27809) doi:10.2196/27809

KEYWORDS

smartphone; hearing aids; app; perception; survey; hearing loss; mobile phone

Introduction

Hearing loss is one of the most common health care problems worldwide. When the World Health Organization started reporting hearing loss in 1985, the number of people with moderate-to-profound hearing impairment was estimated to be 42 million. Furthermore, the number of people with disabling hearing loss reached 466 million in 2018 and is projected to reach approximately 630 million by 2030 [1]. Hearing aids (HAs) are standard hearing intervention methods [2], and the adequate use of HAs improves hearing-specific and general health–related quality of life in adults with mild to moderate hearing loss [3].

Nevertheless, HA adoption rates are extremely low. Globally, only 17% of those who need appropriate hearing rehabilitation use HAs [4]. In addition, a large South Korean cohort study reported that among participants who had minimal hearing loss (mild bilateral hearing loss, unilateral hearing loss, and high frequency hearing loss), only 0.47% of those with subjective symptoms used HAs [5]. Failure to achieve early rehabilitation can accelerate the development of hearing loss and, ultimately, incur enormous social costs [6]. The price of HAs is an important barrier to use [7,8]. When hearing health care is subsidized by the government, HA penetration rates slightly increase [9]. Therefore, other factors, such as social stigma, denial of hearing loss, reduced self-efficacy, and limited access to hearing services should be considered [8]. A prolonged time from the onset of hearing loss to HA intervention has negative effects on quality of life [10]. To address these barriers, alternatives such as over-the-counter (OTC) HAs, personal sound amplification products (PSAPs), and smartphone-based HA apps (SHAAs) have been previously evaluated [11-14]. Moreover, the US Food and Drug Administration announced in its 2016 nonbinding guidance document that medical assessment is no longer required for OTC HAs for individuals aged ≥18 years [15].

SHAAs were originally developed to mimic conventional HA devices. SHAAs refer only to the software installed on a mobile device for hearing support, which is different from the traditional HA hardware–software complex. SHAAs require wired or Bluetooth headsets or headphones instead of hardware resources. Many free or low-price HA apps are available on the web. Although they enhance hearing capabilities through sound amplification, SHAAs were previously far less sophisticated because they could not exactly fit an individual's prescribed target gain as could HAs fitted using real-ear measurement [16]. Some SHAAs have separate channels and advanced functions, such as noise reduction and acoustic feedback suppression [17]. Until now, it was not clear whether SHAAs were clinically effective and could be an alternative device to traditional HAs [18]. In addition, their level of patient satisfaction is generally lower than that with conventional HAs [19]. However, the performance of SHAAs is likely to improve with the development of smartphone hardware and apps, and SHAAs have great potential to contribute to hearing rehabilitation [20].

Easy accessibility is a notable advantage of SHAAs. Users can simply download the app on their smartphones and prepare headsets or headphones for use. SHAAs may particularly help overcome psychological resistance to the use of HAs. Trials with SHAAs showed a reduction in the degree of anxiety and personal distress and increased self-esteem. In addition, reduced stigma or body image of HA users can be expected because of the growing number of individuals who wear headphones with their smartphones [21]. Maidment et al [22] demonstrated that the use of smartphone-connected listening devices in adults with hearing loss could address issues surrounding stigma because smartphones are ubiquitous in everyday life. In addition, the price of SHAAs is lower than that of conventional HAs or PSAPs [23]. Dozens of SHAAs have been released in the App Store (iPhone operating system) and Google Play (Android). Moreover, a new SHAA called Sound Amplifier was introduced by Google [24,25].

The mobile app market is rapidly growing. As of 2019, about 61% of the global population was able to access the internet from mobile devices, and this number is projected to increase to approximately 79% by 2025 [26]. Furthermore, as an increasing number of older adults (>65 years) are using mobile internet via their smartphones, smartphones are expected to exert a greater influence on hearing health care, and SHAAs will expand accordingly [20,27]. Nevertheless, no previous studies have focused on how SHAAs are perceived and the factors affecting the perception of SHAAs. Thus, in this study, we assessed the current awareness of SHAAs and analyzed the associated factors through questionnaires. This information will serve as a baseline for further research on hearing rehabilitation using SHAAs.

Methods

Participants

We performed a multicenter survey of 5 general hospital outpatients who use smartphones and hearing specialists, including otology specialists, audiologists, and HA researchers. Before gaining access to a questionnaire, the potential participants were informed about the survey, and those who

XSL•FO
RenderX

agreed to participate were asked to fill out the questionnaire under the direction of a health care provider.

This study was carried out in accordance with the Declaration of Helsinki on biomedical research for human participants, and the study protocol was approved by the institutional review board of each participating hospital (Seoul St Mary's Hospital, KC20QIDI0526; Chungnam National University Hospital, 2020-06-092; Korea University Hospital, 2020GR0020; Samsung Medical Center, 2020-05-056; and Seoul National University Hospital, D-2003-028-1109).

Questionnaire

A survey on the perception of HAs published by Park et al [28] was modified for use in this study because there is no standardized questionnaire available to assess perceptions of HAs, including SHAAs. Park et al developed a questionnaire that contained 19 questions with an appropriate level of reliability and validity (Cronbach α=.76). To evaluate the consistency of questionnaire items, Hotelling T-square test was used. The items had significant reliability, with F=28.5, P<.001 [28]. One question ("I know that different types of HAs can be worn depending on the degree of hearing loss") was excluded from the questionnaire by Park et al because it was not suitable for the SHAA questionnaires. In addition, *hearing aids* was replaced with *smartphone-based hearing aid apps*. A total of 18 questions in the questionnaire were reviewed by 42 hearing rehabilitation specialists who participated in the opinion survey, and Cronbach α for each question was recalculated (Multimedia Appendix 1). The language used in the questionnaire was Korean. To prevent any possible confusion, respondents were fully informed that SHAAs are independent substitutes for HAs and do not require an additional device other than a smartphone and headphone or headset.

The questionnaire was divided into three sections: (1) sociodemographic characteristics, including age, gender, residence, educational background, economic status, and occupation; (2) clinical characteristics, including the recognition of hearing loss and inconvenience level, the presence of tinnitus and inconvenience level, previous experience with PSAPs or SHAAs by the respondent or their family member, respondent's willingness to use PSAPs or SHAAs, and expected cost of the app; and (3) perception status. In the clinical characteristics section, respondents with hearing loss or tinnitus were asked to assess the degree of their symptoms using a visual analogue scale (VAS). In the perception status section, they were asked to rate 18 questions in 5 categories on a scale from 1 (strongly disagree) to 5 (strongly agree); lower scores indicate poorer awareness. The questions were grouped into 5 categories by similar objectives, which were reviewed by the hearing specialists, allowing the analysis to be simpler and clearer. Questions 1-4 were grouped in the knowledge category, which aimed to evaluate whether respondents were aware of SHAAs as hearing rehabilitation options and how they differed from conventional HAs. Questions 5-6 in the needs category were designed to evaluate whether respondents thought that SHAAs were necessary for hearing discomfort. Questions 7-9 in the cost category were used to identify the influence of price on the decision to purchase, and questions 10-13 were used to evaluate respondents' expectations regarding the ability of SHAAs to improve hearing capabilities. Finally, questions 14-18 in the information category attempted to determine whether participants had accurate information about how to use SHAAs (Multimedia Appendix 1).

The opinion survey for hearing specialists contained questions to determine demographic information such as employment history, educational background, and professional experience (length of career) as well as the expected annual subscription rate for an SHAA and the main selection criteria for HA devices (Multimedia Appendix 2).

Korean Won was used as the standard currency in the questionnaire and was converted into US $ in this report (US $1=₩1082.50).

Statistical Analysis

Age, gender, education background, and economic status data were treated as categorical variables. Reference variables were 20-39 years for age, male for gender, middle school graduate for educational background, and 1 for economic status. VAS score of hearing loss and VAS score of tinnitus were regarded as continuous variables. Linear regression models with the perception level and the expected annual subscription rate as response variables were applied. Robust variance estimation was used for SEs and CIs. Age, gender, educational background, economic status, VAS score of hearing loss, and VAS score of tinnitus were used as explanatory variables in the regression models. Bonferroni-corrected P values <.05 were considered statistically significant. The 2-sample 2-tailed t test was used to compare the expected costs between the hearing specialists and potential users. All statistical analyses were performed using R version 3.6.0 (R Foundation for Statistical Computing).

Results

Clinical Characteristics of Enrolled Participants

A total of 98.6% (219/222) of respondents' answers were analyzed after the survey responses of 3 participants with a survey completion rate <50% were excluded.

The clinical characteristics of the participants are presented in Table 1. The mean age of participants was 52.02 (SD 15.44) years, and male respondents slightly outnumbered female respondents at 59.8% (131/219 male respondents) to 40.2% (88/219 female respondents). Most respondents (138/219, 63%) were college graduates or higher education, and about half (114/219, 52.1%) of the participants estimated themselves as having an intermediate economic status. A total of 44.3% (97/219) of respondents answered that they had subjective hearing loss. The average VAS score of respondents with hearing loss was 2.51 (SD 3.29). In addition, 40.2% (88/219) of respondents had tinnitus, and their average VAS score was 2.39 (SD 3.32). Owing to the multicenter nature of the study, the locations of the participants' residences varied widely. Most participants lived in urban areas (130/219, 59.3%), followed by suburban areas (59/219, 26.9%) and rural areas (30/219, 13.7%). Only 0.9% (2/219) of the enrolled participants had been using HAs at the time of the survey, so wearing HAs was not used in the analysis.

XSL•FO
RenderX

When asked about SHAAs, 21.5% (47/219) of respondents stated that they had considered using an assistive device for hearing loss, but only 3.7% (8/219) respondents knew the difference between traditional HAs and SHAAs. Only 0.9% (2/219) of respondents had experience with an SHAA. However, 26.5% (58/219) of respondents expressed a willingness to use an SHAA in the future.

A total of 42 responses were received from the hearing specialist group, which comprised 29 (69%) otologists and 13 (31%) audiologists. The average number of years working in this profession was 11.83 (SD 7.81) years. A total of 40% (17/42) of respondents had a bachelor's degree, followed by 33% (14/42) of respondents with a master's degree and 26% (11/42) of respondents with a doctorate degree.

Table 1. Participant characteristics (N=219).

Characteristics	Value
Age (years), mean (SD)	52.0 (15.4)
Sex, n (%)	
Male	131 (59.8)
Female	88 (40.2)
Education level, n (%)	
Junior high graduate or less	29 (13.2)
High school graduate	53 (24.2)
College graduate or higher	137 (62.6)
Economic status, n (%)	
A (very low)	12 (5.5)
B (low)	29 (13.2)
C (middle)	113 (51.6)
D (high)	52 (23.7)
E (very high)	13 (5.9)
Subjective hearing loss, n (%)	
Yes	97 (44.3)
No	122 (55.7)
If hearing loss "yes," VAS[a] score (1-10)[b], mean (SD)	2.5 (3.3)
Tinnitus, n (%)	
Yes	88 (40.2)
No	131 (59.8)
If tinnitus "yes," VAS score (1-10)[c], mean (SD)	2.4 (3.3)

[a]VAS: visual analogue scale.

[b]Visual analogue scale (VAS) 1=very minimal problem; VAS 10=very serious problem. VAS 0 was considered to indicate no subjective hearing loss.

[c]Visual analogue scale (VAS) 1=very minimal problem; VAS 10=very serious problem. VAS 0 was considered no subjective tinnitus.

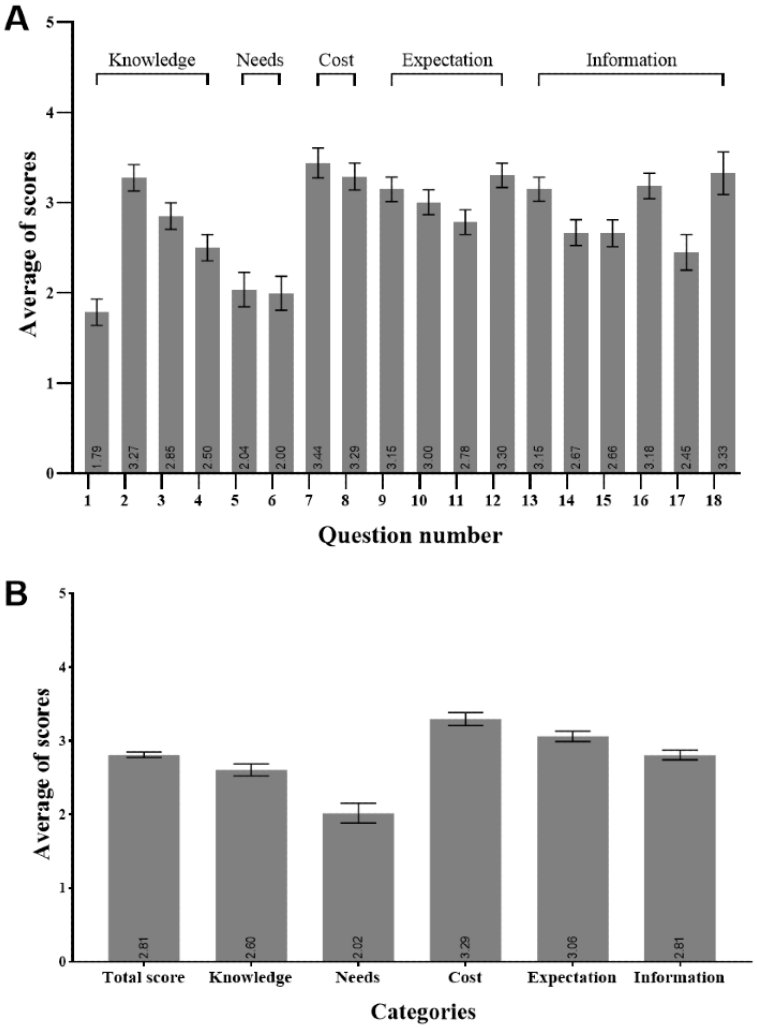
Overall Awareness of SHAAs

The overall score of awareness of SHAAs of the 219 respondents was 2.81 (SD 1.21). Among the 5 categories, the needs category received the lowest score of 2.02 (SD 1.42), whereas the cost category ranked first with a score of 3.29 (SD 1.14; Figure 1).

In the opinion survey of hearing specialists, the main consideration factor for recommending an SHAA was basic performance (30/42, 71%), followed by price (8/42, 19%) and additional functions (2/42, 5%). In addition, noise reduction and the number of channels were mentioned by 1 respondent each.

XSL•FO
RenderX

Figure 1. Average scores obtained in the perception survey. (A) Average scores of 18 questions (rated by question). (B) Average scores of 5 categories (rated by category). Error bars indicate 95% CIs.



Factors Affecting Awareness Scores

Age, Gender, and Area of Residence

Respondents were divided into three age groups: 20-39 years, 40-59 years, and ≥60 years. Compared with the reference age group of 20-39 years, there was no association between age and perception scores in the 40-59–year group, although there was a marginally significant positive correlation between age and information score (P=.05) in this age group. Meanwhile, in the ≥60-year group, there was a remarkable positive correlation between the total SHAA perception score and age in comparison with the 20-39–year group (P=.002), and there were also strong associations between the information and perception scores among the 5 categories (P<.001; Table 2).

To analyze whether gender affected SHAA perception, male respondents were used as the reference group. There were no significant correlations between gender and SHAA perception (Table 3). The area of residence was also not significantly correlated with SHAA perception.

XSL•FO
RenderX

Table 2. Relationship between age and perception scores on smartphone-based hearing aid apps. The reference age group was the 20-39–year group.

Response	Coefficient (SE; 95% CI)	*P* value	Adjusted *P* value
40-59–year group	0.089 (0.095; −0.097 to 0.276)	.35	N/A[a]
Knowledge	−0.155 (0.131; −0.411 to 0.102)	.24	.99
Needs	0.209 (0.155; −0.094 to 0.511)	.18	.89
Cost	0.007 (0.168; −0.323 to 0.336)	.97	.99
Expectation	−0.003 (0.148; −0.293 to 0.288)	.99	.99
Information	0.301 (0.118; 0.069 to 0.533)	.01[b]	.05
≥60-year group	0.314 (0.102; 0.113 to 0.514)	.002[b]	N/A
Knowledge	0.010 (0.151; −0.286 to 0.307)	.95	.99
Needs	0.386 (0.213; −0.031 to 0.803)	.07	.35
Cost	0.286 (0.214; −0.134 to 0.705)	.18	.91
Expectation	0.219 (0.160; −0.095 to 0.533)	.17	.86
Information	0.563 (0.130; 0.308 to 0.819)	<.001[c]	<.001[c]

[a]N/A: not applicable.
[b]*P*<.05.
[c]*P*<.001.

Table 3. Relationship between gender and perception scores regarding smartphone-based hearing aid apps. The reference group was male respondents.

	Coefficient (SE; 95% CI)	*P* value	Adjusted *P* value
Knowledge	−0.053 (0.102; −0.253 to 0.147)	.61	.99
Needs	0.160 (0.170; −0.173 to 0.494)	.35	.99
Cost	0.129 (0.137; −0.140 to 0.397)	.35	.99
Expectation	−0.189 (0.120; −0.424 to 0.046)	.12	.58
Information	−0.027 (0.103; −0.229 to 0.175)	.79	.99
Total	−0.031 (0.079; −0.186 to 0.124)	.70	N/A[a]

[a]N/A: not applicable.

Hearing Loss and Tinnitus

We next evaluated whether subjective hearing loss or tinnitus influenced the perception of SHAA. There were significant correlations between hearing loss and the total perception score (*P*=.001). The presence of hearing loss was strongly associated with the needs category (*P*<.001), but there were no significant associations with the other categories. The degree of hearing loss indicated by the VAS score was closely related to the total scores (*P*=.001) and needs (*P*<.001; Table 4).

Although the presence of tinnitus did not show a significant association with total scores, it was positively correlated with the needs category (*P*=.003). The VAS score for tinnitus did have significant associations with SHAA perception (Table 5).

XSL•FO
RenderX

Table 4. Relationship between subjective hearing loss and perception scores regarding smartphone-based hearing aid apps.

Response	Coefficient (SE; 95% CI)	*P* value	Adjusted *P* value
Subjective hearing loss (yes or no)	0.049 (0.014; 0.021 to 0.076)	.001[a]	N/A[b]
Knowledge	0.021 (0.023; −0.024 to 0.066)	.36	.99
Needs	0.266 (0.030; 0.208 to 0.324)	<.001[c]	<.001[c]
Cost	0.030 (0.027; −0.024 to 0.083)	.28	.99
Expectation	0.024 (0.021; −0.018 to 0.066)	.26	.99
Information	0.017 (0.015; −0.012 to 0.046)	.24	.99
Visual analogue scale score of hearing loss (if hearing loss present)	0.079 (0.024; 0.031 to 0.126)	.001[a]	N/A
Knowledge	0.045 (0.037; −0.028 to 0.117)	.23	.99
Needs	0.304 (0.066; 0.175 to 0.434)	<.001[c]	<.001[c]
Cost	0.024 (0.040; −0.054 to 0.102)	.55	.99
Expectation	0.059 (0.034: −0.008 to 0.127)	.08	.42
Information	0.057 (0.026; 0.006 to 0.107)	.03[d]	.14

[a]*P*<.01.
[b]N/A: not applicable.
[c]*P*<.001.
[d]*P*<.05.

Table 5. Relationship between subjective tinnitus and perception scores regarding smartphone-based hearing aid apps.

Response	Coefficient (SE; 95% CI)	*P* value	Adjusted *P* value
Subjective tinnitus (yes or no)	0.003 (0.012; −0.022 to 0.027)	.82	N/A[a]
Knowledge	0.049 (0.020; 0.011 to 0.088)	.01[b]	.06
Needs	−0.094 (0.027; −0.147 to −0.041)	.001[c]	.003[b]
Cost	−0.001 (0.024; −0.047 to 0.046)	.98	.99
Expectation	0.022 (0.020; −0.018 to 0.061)	.28	.99
Information	−0.007 (0.014; −0.034 to 0.020)	.60	.99
Visual analogue scale score of tinnitus (if hearing loss present)	0.021 (0.022; −0.022 to 0.064)	.34	N/A
Knowledge	0.068 (0.031; 0.007 to 0.128)	.03[b]	.14
Needs	−0.089 (0.056; −0.200 to 0.021)	.11	.57
Cost	0.049 (0.044; −0.037 to 0.135)	.26	.99
Expectation	0.052 (0.035; −0.018 to 0.121)	.15	.72
Information	−0.003 (0.026; −0.054 to 0.049)	.92	.99

[a]N/A: not applicable.
[b]*P*<.05.
[c]*P*<.01.

Expected Price of SHAAs

The average expected cost for an annual subscription to an SHAA was US $84.43 (95% CI 75.66-93.21). Analyzed by age group, the average expected prices were US $97.37 (95% CI 75.10-119.54), US $78.98 (95% CI 63.46-94.41), and US $86.47 (95% CI 62.91-109.93) in the 20-39–year group, the 40-59–year group, and in the ≥60-year group, respectively (Figure 2). The expected cost was significantly correlated with economic status (*P*=.02), whereas it was not significantly associated with other categories (Table 6).

The experts' average expected cost for an annual subscription to a premium version app was US $32.48 (95% CI 17.81-47.24), and 33% (14/42) of respondents answered that the app should be available at no cost. As for an entry-version app, the expected cost for an annual subscription was US $9.69 (95% CI

XSL•FO
RenderX

2.68-16.70) on average, and 71% (30/42) of respondents expected this app to be provided free of charge.

The average cost for an annual subscription expected by potential users was markedly higher than that expected by hearing specialists (based on the premium version; *P*<.001). A total of 45% (19/42) of respondents among the potential users were not willing to pay for the app, which was much lower than the percentage of the hearing specialist group who thought the SHAA should be provided for free (*P*<.001).

Figure 2. Average expected cost for smartphone-based hearing aid apps according to age group. Error bars indicate 95% CIs.

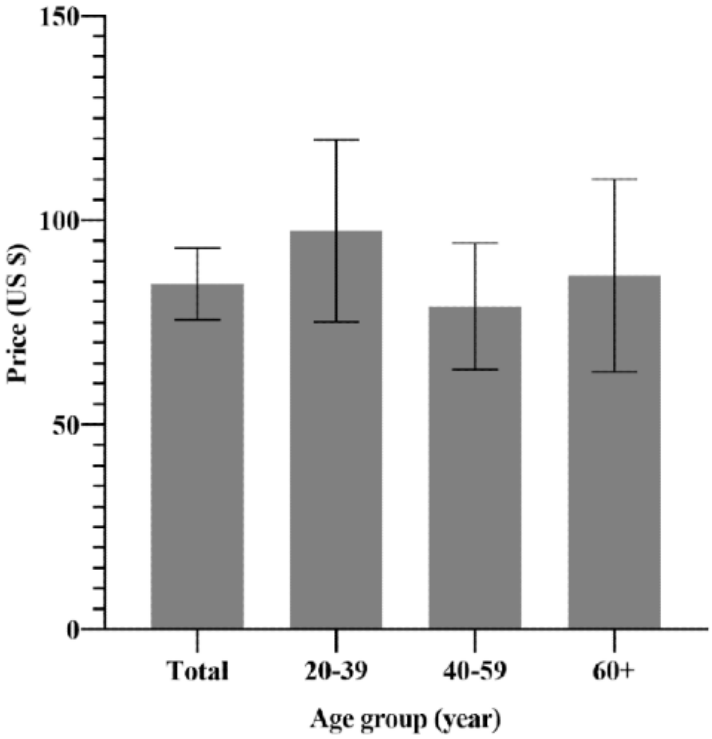


Table 6. Factors affecting the expected price of smartphone-based hearing aid apps.

Variable	Coefficient (SE; 95% CI)	*P* value
Age (years)		
40-59	−1.991 (1.348; −4.633 to 0.651)	.14
≥60	−1.180 (1.509; −4.137 to 1.778)	.43
Sex (female)	0.971 (0.983; −0.955 to 2.898)	.32
Education level		
High school graduate	−1.156 (1.875; −4.831 to 2.518)	.54
University graduate	0.122 (1.997; −3.793 to 4.037)	.95
Economic status	−1.474 (0.638; −2.723 to −0.224)	.02[a]
Hearing loss visual analogue scale score	0.261 (0.209; −0.148 to 0.671)	.21
Tinnitus visual analogue scale score	−0.207 (0.179; −0.559 to 0.145)	.25

[a]*P*<.05.

XSL•FO
RenderX

Discussion

Principal Findings

This cross-sectional study recruited 261 participants, consisting of 219 outpatients and 42 hearing specialists from multiple locations in South Korea, to avoid regional bias. In addition, as non–smartphone users are not potential candidates for SHAA use, this study targeted people who own and use smartphones. Overall, only a limited number of participants had heard about SHAAs, and only 4% (9/219) of respondents were aware of the differences between SHAAs and conventional HAs. In addition, only 0.9% (2/219) of the respondents had experience using an SHAA. These results indicated an extremely low level of perception regarding SHAAs. However, it is noteworthy that 26% (57/219) of the respondents stated their intention to consider using SHAAs after they obtained information about SHAAs during the survey. This suggests that increased awareness of SHAAs may lead to their use by more individuals with hearing loss.

As the perception of SHAAs was more meaningful to people with hearing problems than to the general population, the survey was conducted for ear, nose, and throat outpatients and resulted in a relatively high proportion of participants with hearing loss or tinnitus. There was also a strong association between the perception of SHAAs and the age and degree of hearing loss. The amount of information increased with age, whereas gender showed no relationship with the information. Although the degree of hearing loss influenced the purchase of an SHAA, the level of tinnitus was not related to the perception of SHAAs. These findings suggest that the demand for hearing rehabilitation devices increases with age and the development of hearing loss, indicating the necessity of providing further relevant information to elderly individuals with hearing impairment.

The expected cost was associated with the economic status. The prices that respondents were willing to pay for SHAA were relatively high in the 20-39–year group and the >60-year group. This is perhaps because the younger generation group would like to improve their own or their parents' hearing capacity, and the older adult group faces more inconvenience from hearing loss. In addition, price was regarded as one of the most crucial factors determining whether or not to purchase an SHAA. Respondents anticipated more advanced features with an increase in price. The average expected cost for an annual subscription to an SHAA was higher than that expected by the hearing specialists. We assumed that those with hearing impairment were willing to pay a higher price than expected by specialists because of the effect of hearing loss on their quality of life. Furthermore, the expected cost was higher than the actual price of Petralex (once-off annual cost: US $59.99 for iPhone operating system), one of the most expensive SHAAs on the market [20,29]. It is notable that only 5.9% (13/219) of respondents expected the app to be free. This suggests that potential users are willing to pay a certain amount for an SHAA with the expectation of efficacy. Nevertheless, the expected cost is substantially lower than that of commercially available HAs or PSAPs. The cost of HA fitting for a single device was US $2336 in the United States [30,31]. Moreover, OTC HAs range in price from approximately US $600 to US $1000 [32], and lower-priced PSAPs range from US $250 to US $350 [33]. The life expectancy of HAs or PSAPs is approximately 5 years. A 5-year subscription to an SHAA would be approximately US $430, which is much lower than the price of HAs and similar to that of premium PSAPs. Thus, SHAAs are likely to compete with PSAPs for market share in the future.

Smartphone-based mobile health is widely used for diagnostics and therapy [34] and also supports hearing rehabilitation. Paglialonga et al [25] investigated 200 hearing health care apps available on the market. Among these apps, the largest proportion (28%) comprised sound enhancement apps [25].

SHAAs have several advantages. First, SHAAs range in price from free to US $70, and are therefore cheaper than conventional HAs overall. SHAAs are therefore likely to substitute for traditional HAs [9]. Second, patients with hearing loss can receive a call and perform HA fitting directly with their smartphones [35]. Third, because of the convenience offered by smartphones in our daily lives, SHAAs may allow patients with hearing loss to feel free from the stigma of using HAs [22]. Finally, the advantages mentioned enable SHAAs to act as *gateway products* to more sophisticated devices, such as conventional HAs [36].

South Korea's gross domestic product per capita is US $32,310, ranking South Korea 28th across the globe [37]. In particular, South Korea has one of the highest smartphone penetration rates, with the smallest gap among all ages (percentage of adults who own a smartphone in South Korea in 2018: 18-34–year group, 99% and >50-year group, 91%) [38]. Given that smartphone use is skyrocketing worldwide, awareness of SHAAs can increase global accessibility to HA interventions.

However, it should be noted that the effectiveness of SHAAs has not been fully proven. Amlani et al [23] recommended that SHAAs be used only as a temporary means of assistance by patients using HA. Medwetsky et al [39] reported that SHAAs improved listening performance, but test participants had only mild to moderate high frequency hearing loss. As the effectiveness of SHAAs in patients with moderate to severe hearing loss is yet to be determined, it is essential to carry out a series of well-designed studies to determine the efficacy of SHAAs in hearing rehabilitation.

Comparison With Previous Work

Previous studies have demonstrated that SHAAs can improve hearing performance in patients with and without hearing loss [20,40]. Most previous studies compared auditory performance with conventional HAs in patients with hearing loss. They evaluated the self-reported benefits and satisfaction in a small case series in a single center [20,21]. Performance and satisfaction show wide variations according to app, operation system, and type of headphones [18]. However, previous studies did not comprehensively evaluate the awareness and associated factors of SHAAs in a cross-sectional multicenter survey. Most of the participants in our study were non-HA users. As less than half of them had subjective hearing loss, we think they could be potential candidates for SHAAs. In addition, this study showed the expectation and expected cost of HAs in smartphone

XSL•FO
RenderX

users and hearing professionals. Our data showed that the price of SHAAs is underestimated and suggested an expected cost, which is useful information for mobile app users and developers.

Strengths and Limitations

One strength of this study is that it is the first study to measure the perception of SHAAs. These findings are expected to pave the way for more surveys regarding awareness of other hearing rehabilitation devices. In addition, because of the multicenter nature of this study, our findings are generalizable to a broad population. According to a survey conducted by the government, the urban population of South Korea is approximately 90% of the total population, which is similar to the population distribution in our results (189/219, 86.3%) [41].

One limitation of this study is that because no standardized questionnaires are available to evaluate perceptions of hearing rehabilitation devices, we modified the questionnaire of a preceding study that investigated awareness about HAs [28]. As this questionnaire was not originally designed or validated to measure perceptions of SHAAs, our findings should be interpreted with caution. A well-validated survey on the perception of hearing assistant devices such as SHAAs should be developed in the future.

In addition, the participants in this study were younger than those of known typical HA seekers [10]. In addition, we did not investigate the experience of HAs because of the small number of HA users in this study. HA users may not actively seek alternative devices, and these HA users could have altered the results of the survey.

Furthermore, our findings do not provide insight into the efficacy of SHAAs in remediating hearing loss; our focus was primarily on the perception of SHAAs. Thus, clinical validation of the effectiveness of the SHAA is required [21].

Conclusions

SHAAs are an alternative hearing rehabilitation option for smartphone users with hearing loss who have no access to appropriate hearing rehabilitation devices because of their high costs. However, the perception of SHAAs was very low. Age and degree of hearing loss were correlated with perception scores. Potential users estimated the cost of an SHAA as approximately US $86 for a 1-year subscription. Those with hearing loss and requiring hearing rehabilitation were willing to pay a higher price than what the hearing specialists expected the price to be. In addition, a higher economic status was associated with an increased willingness to pay higher prices. Considering that a large portion of respondents showed interest in SHAA after obtaining information from the survey, enhancement of perception of SHAAs is likely crucial to expand their market base.

Acknowledgments

This research was supported by a grant from the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant HC19C0128).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey on the perception of hearing aid apps.
[PDF File (Adobe PDF File), 322 KB - mhealth_v10i1e27809_app1.pdf]

Multimedia Appendix 2
Hearing specialist's opinion for hearing aid apps.
[PDF File (Adobe PDF File), 207 KB - mhealth_v10i1e27809_app2.pdf]

References

1. Davis AC, Hoffman HJ. Hearing loss: rising prevalence and impact. Bull. World Health Organ 2019 Oct 01;97(10):646-646A. [doi: 10.2471/blt.19.224683]
2. Chisolm TH, Johnson CE, Danhauer JL, Portz LJ, Abrams HB, Lesner S, et al. A systematic review of health-related quality of life and hearing aids: final report of the American Academy of Audiology Task Force on the health-related quality of life benefits of amplification in adults. J Am Acad Audiol 2007 Feb 07;18(2):151-183. [doi: 10.3766/jaaa.18.2.7] [Medline: 17402301]
3. Ferguson MA, Kitterick PT, Chong LY, Edmondson-Jones M, Barker F, Hoare DJ. Hearing aids for mild to moderate hearing loss in adults. Cochrane Database Syst Rev 2017 Sep 25;9:CD012023 [FREE Full text] [doi: 10.1002/14651858.CD012023.pub2] [Medline: 28944461]
4. Orji A, Kamenov K, Dirac M, Davis A, Chadha S, Vos T. Global and regional needs, unmet needs and access to hearing aids. Int J Audiol 2020 Mar 03;59(3):166-172. [doi: 10.1080/14992027.2020.1721577] [Medline: 32011190]
5. Choi JE, Ahn J, Park HW, Baek S, Kim S, Moon IJ. Prevalence of minimal hearing loss in South Korea. PLoS One 2017 Feb 14;12(2):e0171635 [FREE Full text] [doi: 10.1371/journal.pone.0171635] [Medline: 28196098]

XSL•FO
RenderX

6. Deafness and hearing loss. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [accessed 2021-12-14]
7. Grundfast KM, Liu SW. What otolaryngologists need to know about hearing aids. JAMA Otolaryngol Head Neck Surg 2017 Feb 01;143(2):109-110. [doi: 10.1001/jamaoto.2016.3416] [Medline: 27893006]
8. McKee M, Choi H, Wilson S, DeJonckheere M, Zazove P, Levy H. Determinants of hearing aid use among older Americans with hearing loss. Gerontologist 2019 Nov 16;59(6):1171-1181 [FREE Full text] [doi: 10.1093/geront/gny051] [Medline: 29788270]
9. Valente M, Amlani AM. Cost as a barrier for hearing aid adoption. JAMA Otolaryngol Head Neck Surg 2017 Jul 01;143(7):647-648. [doi: 10.1001/jamaoto.2017.0245] [Medline: 28520884]
10. Simpson AN, Matthews LJ, Cassarly C, Dubno JR. Time from hearing aid candidacy to hearing aid adoption: a longitudinal cohort study. Ear Hear 2019;40(3):468-476 [FREE Full text] [doi: 10.1097/AUD.0000000000000641] [Medline: 30085938]
11. Chan Z, McPherson B. Over-the-counter hearing aids: a lost decade for change. Biomed Res Int 2015;2015:827463 [FREE Full text] [doi: 10.1155/2015/827463] [Medline: 26557701]
12. Cho YS, Park SY, Seol HY, Lim JH, Cho Y, Hong SH, et al. Clinical performance evaluation of a personal sound amplification product vs a basic hearing aid and a premium hearing aid. JAMA Otolaryngol Head Neck Surg 2019 Jun 01;145(6):516-522 [FREE Full text] [doi: 10.1001/jamaoto.2019.0667] [Medline: 31095263]
13. Choi JE, Kim J, Yoon SH, Hong SH, Moon IJ. A personal sound amplification product compared to a basic hearing aid for speech intelligibility in adults with mild-to-moderate sensorineural hearing loss. J Audiol Otol 2020 Apr;24(2):91-98 [FREE Full text] [doi: 10.7874/jao.2019.00367] [Medline: 31842534]
14. Franklin C. Over the counter hearing aids. J Phonet Audiol 2017;03(01). [doi: 10.4172/2471-9455.1000e106]
15. Immediately in effect guidance document: conditions for sale for air-conduction hearing aids - guidance for industry and food and drug administration staff. Center for Devices and Radiological Health. 2016. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/immediately-effect-guidance-document-conditions-sale-air-conduction-hearing-aids [accessed 2021-12-14]
16. Aazh H, Moore BC. The value of routine real ear measurement of the gain of digital hearing aids. J Am Acad Audiol 2007 Sep 07;18(8):653-664. [doi: 10.3766/jaaa.18.8.3] [Medline: 18326152]
17. Vashkevich M, Azarov E, Petrovsky N, Petrovsky A. Petralex: a smartphone-based real-time digital hearing aid with combined noise reduction and acoustic feedback suppression. In: Proceedings of the Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA) Conference. 2017 Presented at: Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA) Conference; Sept. 20-22, 2017; Poznan, Poland. [doi: 10.23919/spa.2017.8166873]
18. Maidment DW, Amlani AM. Argumentum ad ignorantiam: smartphone-connected listening devices. Semin Hear 2020 Nov 16;41(4):254-265 [FREE Full text] [doi: 10.1055/s-0040-1718711] [Medline: 33364675]
19. Maidment DW, Ferguson M. An application of the Medical Research Council's guidelines for evaluating complex interventions: a usability study assessing smartphone-connected listening devices in adults with hearing loss. Am J Audiol 2018 Nov 19;27(3S):474-481. [doi: 10.1044/2018_aja-imia3-18-0019]
20. De Sousa KC, Moore DR, Motlagh-Zadeh L, Myburgh HC, Swanepoel DW. Do smartphone hearing aid apps work? Hearing J 2019;72(11):34. [doi: 10.1097/01.hj.0000612588.11985.6a]
21. Amlani AM, Smaldino J, Hayes D, Taylor B, Gessling E. Feasibility of using a smartphone-based hearing aid application to improve attitudes toward amplification and hearing impairment. Am J Audiol 2019 Mar 15;28(1):125-136. [doi: 10.1044/2018_aja-17-0068]
22. Maidment DW, Ali YH, Ferguson MA. Applying the COM-B model to assess the usability of smartphone-connected listening devices in adults with hearing loss. J Am Acad Audiol 2019 May;30(5):417-430. [doi: 10.3766/jaaa.18061] [Medline: 31044693]
23. Amlani A, Taylor B, Levy C, Robbins R. Utility of smartphone-based hearing aid applications as a substitute to traditional hearing aids. The Hearing Review. 2013. URL: https://hearingreview.com/hearing-products/utility-of-smartphone-based-hearing-aid-applications-as-a-substitute-to-traditional-hearing-aids [accessed 2021-12-14]
24. Sound amplifier. Google LLC. 2020. URL: https://play.google.com/store/apps/details?id=com.google.android.accessibility.soundamplifier&hl=en_US&gl=US [accessed 2021-12-14]
25. Paglialonga A, Tognola G, Pinciroli F. Apps for hearing healthcare. Stud Health Technol Inform 2015;210:666-668. [Medline: 25991233]
26. The mobile economy 2019. GSM Association. 2019. URL: https://www.gsmaintelligence.com/research/?file=b9a6e6202ee1d5f787cfebb95d3639c5&download [accessed 2021-12-14]
27. Kimball SH, Singh G, John AB, Jenstad LM. Implications and attitudes of audiologists towards smartphone integration in hearing healthcare. Hear Res 2018 Nov;369:15-23. [doi: 10.1016/j.heares.2018.06.011] [Medline: 29945802]
28. Park SG, Yu CM, Chung CH, Kim MK, Huh MJ. A study on the perception of hearing aids according to hearing loss degree for the elderly in Busan. J Speech-Lang Hear Disord 2012 Jun;21(2):227-241. [doi: 10.15724/jslhd.2012.21.2.013]
29. Petralex hearing aid app. Petralex. 2020. URL: https://apps.apple.com/us/app/apple-store/id816133779 [accessed 2021-12-14]

XSL•FO
RenderX

30. Bailey A. How much do hearing aids cost in 2021? Hearing Tracker. 2021. URL: https://www.hearingtracker.com/how-much-do-hearing-aids-cost#a-hearing-aid-prices [accessed 2021-12-14]
31. Lin FR, Hazzard WR, Blazer DG. Priorities for improving hearing health care for adults: a report from the National Academies of Sciences, Engineering, and Medicine. J Am Med Assoc 2016 Aug 23;316(8):819-820. [doi: 10.1001/jama.2016.7916] [Medline: 27254725]
32. Akers W. Prescription hearing aids are expensive. Could an OTC option help? Healthline. 2020. URL: https://www.healthline.com/health-news/why-is-it-so-hard-to-get-a-hearing-aid-otc [accessed 2021-12-14]
33. Gandel C. Not ready for a hearing aid but need a little help? AARP. 2015. URL: https://www.aarp.org/health/conditions-treatments/info-2015/hearing-amplifiers-psaps.html [accessed 2021-12-14]
34. Sim I. Mobile devices and health. N Engl J Med 2019 Sep 05;381(10):956-968. [doi: 10.1056/nejmra1806949]
35. Aldaz G, Puria S, Leifer LJ. Smartphone-based system for learning and inferring hearing aid settings. J Am Acad Audiol 2016 Oct 06;27(9):732-749 [FREE Full text] [doi: 10.3766/jaaa.15099] [Medline: 27718350]
36. Olson A, Maidment DW, Ferguson MA. Consensus on connected hearing health technologies and service delivery models in the UK: a Delphi review. Int J Audiol 2021 Jun 28:1-8. [doi: 10.1080/14992027.2021.1936223] [Medline: 34182863]
37. GDP per capita, current prices. International Monetary Fund. 2021. URL: https://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEOWORLD [accessed 2021-12-14]
38. Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2021-12-14]
39. Medwetsky L, Kelly K, Bakke M. Earphone models for iPhones: surprising results when used with a hearing app. The Hearing Review. 2020. URL: https://hearingreview.com/hearing-products/hearing-aids/apps/earphone-models-for-iphones-surprising-results-when-used-with-a-hearing-app [accessed 2020-09-24]
40. Martinez-Beneyto P, Franchella S, Rodriguez FA, Navarro-Velasquez R, Martinez-Beneito MA, Martini A, et al. Are smartphone applications (App) useful to improve hearing? Acta Otorhinolaryngol Ital 2020 Aug;40(4):304-310 [FREE Full text] [doi: 10.14639/0392-100X-N0318] [Medline: 32519992]
41. Trend status of population ratio in urban areas. Ministry of Land, Infrastructure and Transport. 2020. URL: https://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1200 [accessed 2021-12-14]

Abbreviations

HA: hearing aid
OTC: over-the-counter
PSAP: personal sound amplification product
SHAA: smartphone-based hearing aid app
VAS: visual analogue scale

Edited by L Buis; submitted 08.02.21; peer-reviewed by Y Oh, J Jung, D Maidment; comments to author 23.03.21; revised version received 06.06.21; accepted 24.11.21; published 07.01.22.

Please cite as:
Han JS, Park YH, Song JJ, Moon IJ, Lee W, Kim Y, Cho YS, Seo JH, Park MK
Knowledge and Expectations of Hearing Aid Apps Among Smartphone Users and Hearing Professionals: Cross-sectional Survey
JMIR Mhealth Uhealth 2022;10(1):e27809
URL: https://mhealth.jmir.org/2022/1/e27809
doi:10.2196/27809
PMID:34994699

©Jae Sang Han, Yong-Ho Park, Jae-Jun Song, Il Joon Moon, Woojoo Lee, Yoonjoong Kim, Young Sang Cho, Jae-Hyun Seo, Moo Kyun Park. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 07.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Applying an Extended UTAUT2 Model to Explain User Acceptance of Lifestyle and Therapy Mobile Health Apps: Survey Study

Eva-Maria Schomakers[1], MSc; Chantal Lidynia[1], PhD; Luisa Sophie Vervier[1], DiplPsych; André Calero Valdez[1], PhD; Martina Ziefle[1], PhD, Prof Dr

Human-Computer Interaction Center, RWTH Aachen University, Aachen, Germany

Corresponding Author:
Eva-Maria Schomakers, MSc
Human-Computer Interaction Center
RWTH Aachen University
Campus-Boulevard 57
Aachen, 52074
Germany
Phone: 49 17662315488
Email: schomakers@comm.rwth-aachen.de

Abstract

Background: Mobile health (mHealth) care apps are a promising technology to monitor and control health individually and cost-effectively with a technology that is widely used, affordable, and ubiquitous in many people's lives. Download statistics show that *lifestyle apps* are widely used by young and healthy users to improve fitness, nutrition, and more. While this is an important aspect for the prevention of future chronic diseases, the burdened health care systems worldwide may directly profit from the use of *therapy apps* by those patients already in need of medical treatment and monitoring.

Objective: We aimed to compare the factors influencing the acceptance of lifestyle and therapy apps to better understand what drives and hinders the use of mHealth apps.

Methods: We applied the established unified theory of acceptance and use of technology 2 (UTAUT2) technology acceptance model to evaluate mHealth apps via an online questionnaire with 707 German participants. Moreover, trust and privacy concerns were added to the model and, in a between-subject study design, the influence of these predictors on behavioral intention to use apps was compared between lifestyle and therapy apps.

Results: The results show that the model only weakly predicted the intention to use mHealth apps (R^2=0.019). Only hedonic motivation was a significant predictor of behavioral intentions regarding both app types, as determined by path coefficients of the model (lifestyle: 0.196, P=.004; therapy: 0.344, P<.001). Habit influenced the behavioral intention to use lifestyle apps (0.272, P<.001), while social influence (0.185, P<.001) and trust (0.273, P<.001) predicted the intention to use therapy apps. A further exploratory correlation analysis of the relationship between user factors on behavioral intention was calculated. Health app familiarity showed the strongest correlation to the intention to use (r=0.469, P<.001), stressing the importance of experience. Also, age (r=–0.15, P=.004), gender (r=–0.075, P=.048), education level (r=0.088, P=.02), app familiarity (r=0.142, P=.007), digital health literacy (r=0.215, P<.001), privacy disposition (r=–0.194, P>.001), and the propensity to trust apps (r=0.191, P>.001) correlated weakly with behavioral intention to use mHealth apps.

Conclusions: The results indicate that, rather than by utilitarian factors like usefulness, mHealth app acceptance is influenced by emotional factors like hedonic motivation and partly by habit, social influence, and trust. Overall, the findings give evidence that for the health care context, new and extended acceptance models need to be developed with an integration of user diversity, especially individuals' prior experience with apps and mHealth.

(JMIR Mhealth Uhealth 2022;10(1):e27095) doi:10.2196/27095

KEYWORDS

technology acceptance; UTAUT2; mHealth; privacy concerns; trust

XSL•FO
RenderX

Introduction

Overview

Due to their affordability and ubiquity in people's everyday lives [1], smartphone apps offer the opportunity to monitor and control health more individually and cost-effectively than ever before [2]. Mobile health (mHealth) care apps are seen as a promising technology that has the potential to improve people's health in general; specifically, they can, for example, enhance the independence of chronically ill people [3,4] and improve rehabilitation success [5] or outcomes of diabetes self-management [6].

mHealth apps encompass a variety of health-related services (eg, support of diagnostics and treatment), tracking of infection processes (eg, contact tracing during the COVID-19 pandemic), remote monitoring, and medicine intake reminders [7]. In addition to enhancing quality of life and supporting medical therapy, mHealth apps also support chronic disease prevention (eg, with nutrition and activity monitoring), as a sedentary lifestyle is one of the key problems of our societies that lead to an increase in chronic disease prevalence [8].

Users' technology acceptance is one decisive factor for the adoption and widespread use of technologies, including mHealth apps, but it can also be a barrier if the diverse requirements of the potential users are not understood [9]. Established technology acceptance models, such as the unified theory of acceptance and use of technology 2 (UTAUT2) [10], structure important predictive factors for the intention to use mHealth apps. However, these established models have been criticized as not being fully applicable to the health care context, as relevant factors are missing [11,12].

This research aims to improve the understanding of users' decisions to use, or reject the use of, mHealth apps. To do so, we have built on and extended the established and widely used UTAUT2 acceptance model. As mHealth apps are essentially based on the collection and analysis—and often also transmission—of user data, privacy concerns can be a reason for rejection. Medical data are perceived as very sensitive and, thus, even more reluctantly disclosed [13]. Also, trust or distrust in the reliability and competence of technologies is an important predictor for their use [14]. Still, these factors are missing from established and widely used technology acceptance models, like the UTAUT2. Therefore, we integrated both privacy concerns and trust in the mHealth app into the UTAUT2 model and empirically tested their impact on mHealth app acceptance.

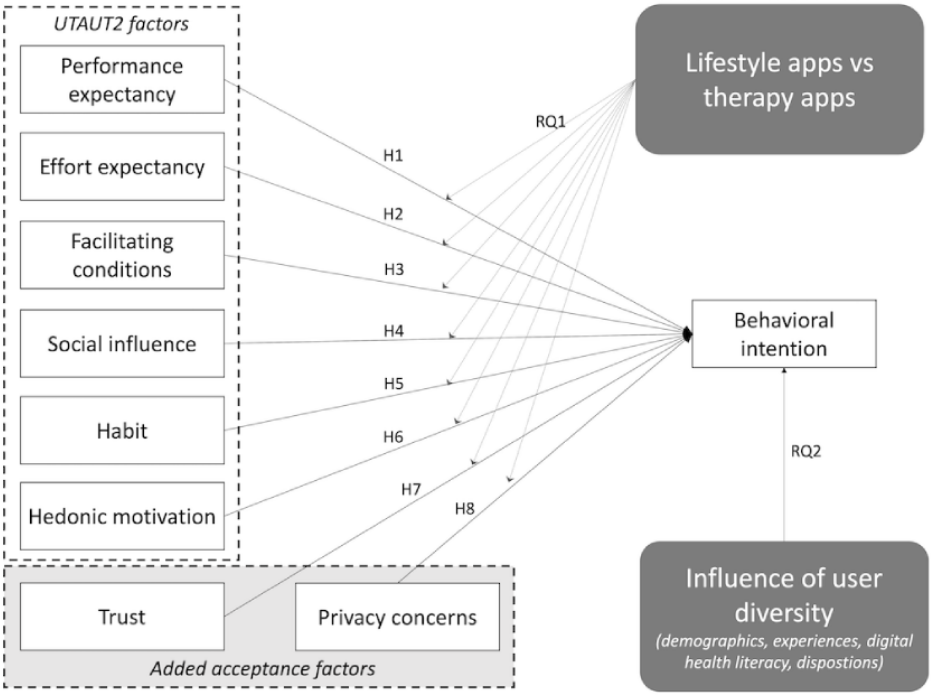
Moreover, we tested for differences in acceptance patterns between two types of mHealth apps (research question 1): currently, therapy apps targeted at existing illnesses and ailments are far less often used than lifestyle apps that, for example, should improve fitness and prevent health problems [15]. Therefore, the question arises as to whether there are differences between the factors shaping the acceptance of therapy apps compared to lifestyle apps. To study these differences, we employed our extended UTAUT2 model to compare the acceptance of therapy and lifestyle apps.

Another important research duty regarding the acceptance of mHealth apps is the integration of effects of user diversity on technology acceptance. While demographic factors, such as age, have already been a focus of research (eg, Deng et al [14], Risch [15], and Guo et al [16]), the influence of experience, digital health literacy, and personal dispositions still needs to be further understood. While we did not extensively study the impact of these factors, we nevertheless considered their importance and exploratorily analyzed their relationship to mHealth acceptance to lay a basis for future research (research question 2).

Figure 1 depicts our research model, based on the UTAUT2 with the inclusion of privacy concerns and trust. The hypotheses and research questions will be developed and explained in the following sections.

XSL•FO
RenderX

Figure 1. Proposed research model. H: hypothesis; RQ: research question; UTAUT2: unified theory of acceptance and use of technology 2.



Our research provides new insights into the individual and context-specific acceptance patterns for mHealth apps. This is important for revealing drivers to build acceptance as well as for identifying barriers that need to be reduced. User acceptance is one of the keys to a successful mHealth app rollout and to harnessing the full potential of mHealth for health care systems and for the improvement of quality of life and therapy for patients.

The UTAUT2 as a Theoretical Framework

With the ever-increasing use of technology, the acceptance of new devices and software has advanced as an important focus of research. To understand the future use of new implementations, it is necessary to understand what factors influence human behavior. Based on psychological theories (eg, the theory of reasoned action [17]), technology acceptance models argue that the actual use of technology is largely influenced by a previous intention to use it, thus justifying the inclusion of not only current users but also potential future users into acceptance studies.

The most recent extension in the line of technology acceptance models is the UTAUT2 [10], which focuses on the acceptance of commercially available technologies. The UTAUT2 is often cited and has been applied and extended a multitude of times in various technology contexts (see recent reviews of the UTAUT2 [18,19]). In addition to performance expectancy, effort expectancy, social influence, and facilitating conditions, which were already part of the predecessor, UTAUT [20], in the context of private use of information technologies, factors such as fun (ie, hedonic motivation), finances (ie, price value), and habit were found to influence the behavioral intention to use technology and, thus, acceptance.

UTAUT2 Constructs

Performance expectancy describes the perception that using the technology will provide benefits to the user and is, thus, tied to the perception of usefulness [10]. A recent review and weight analysis of UTAUT2 studies [19] showed that, indeed, in the large majority of studies, performance expectancy significantly influenced use intention and was also often the strongest predictor. Further, in mHealth research specifically, performance expectancy was continuously shown to have a significant impact on use intentions [21-26]. This applies to studies regarding lifestyle apps (eg, Schomakers et al [21] and Yuan et al [23]), therapy apps (eg, Schomakers et al [21] and Hoque and Sorwar [24]), as well as mHealth apps in general (eg, Sun et al [22] and Salgado et al [26]).

Based on this body of research, hypothesis 1 is as follows: Performance expectancy influences the intention to use mHealth apps.

Effort expectancy describes the expected ease of using the technology [10]. Results regarding the influence of effort expectancy on use intention are mixed [19]. Hoque and Sorwar [24] found effort expectancy to be a significant predictor of mHealth acceptance by older adult users, and Wang et al [25] found it to predict the intention to use online hospital mHealth services in China. Other studies could not confirm an impact on use intention of fitness and diabetic mHealth apps [21,23] or of mHealth services in general [26].

In this study, we, therefore, again examined this relationship using hypothesis 2: Effort expectancy influences the intention to use mHealth apps.

Social influence is “the degree to which an individual perceives that important others believe he or she should use the new system” [20]. Results on the significance of its influence on use intention are mixed, as approximately half of the studies

XSL•FO
RenderX

applying the UTAUT2 found a significant effect [19] of social influence on acceptance. In the mHealth context, Schomakers et al [21] found social influence to be significant for both diabetes apps and fitness apps, but it showed a stronger influence on the use intention for diabetes apps. Regarding the use of mHealth by older adult users, Hoque and Sorwar [24] found a significant impact, but other studies could not confirm the effect in mHealth [23,25,26].

Therefore, hypothesis 3 is as follows: Social influence affects the intention to use mHealth apps.

Facilitating conditions refer to the perceptions "of the resources and support available to perform a behavior" [10]. For the use of mHealth apps, this is regarding, for example, a smart device on which to use apps or peers to ask about problems interacting with apps. Again, previous results are also mixed depending on the technology researched [19]. Regarding mHealth, only Wang et al [25] and Sun et al [22] found significant effects on use intention.

In this regard, we put forward hypothesis 4: Facilitating conditions influence the intention to use mHealth apps.

Hedonic motivation is the gratification counterpart to the utilitarian measure of usefulness represented by performance expectancy. Hedonic motivation refers to fun, pleasure, and enjoyment with the use of technology [10]. It is a rather strong predictor in many studies applying the UTAUT2 model [19]. In mHealth, Yuan et al [23] found a moderate effect on the intention to use fitness apps, and in a qualitative study, Woldeyohannes and Ngwenyama [27] also found evidence for its importance on mHealth app acceptance. However, Salgado et al [26] could not confirm this for the general use of mHealth.

Considering these mixed results, we examined the following relationship as hypothesis 5: Hedonic motivation influences the intention to use mHealth apps.

Habit is operationalized within the UTAUT2 framework as a self-reported perception of a customary use of the respective technology [10]. Castanha et al [19] found in their review of UTAUT2 studies that habit had good predictive abilities for use intention, and almost all studies that included habit confirmed its impact. Salgado et al [26] found habit to be the strongest predictor for mHealth use intention and found it to be the only significant one from the UTAUT2 model besides performance expectancy. The impact of habit was also confirmed for fitness apps [23]. Still, habit only has relevance for those people who already use mHealth apps.

Therefore, in this study, we only assessed habit for current users of the mHealth apps in question and proposed hypothesis 6: Habit influences the intention to use mHealth apps of current users of mHealth apps.

In the UTAUT2, a seventh predictor is the price value. Despite its significant effect on use intention as shown in some UTAUT2 studies [19,23], we did not integrate price value in our model, as most existing and well-known mHealth apps are free of charge. To include the effects of cost or price for such apps in the analysis runs the risk of price or cost obscuring all other acceptance factors, simply because people tend to reject or ascribe less value to things that are currently unaffordable to them [28]. However, we were interested in first identifying the interaction and relationship of the other acceptance factors. Therefore, the topic of cost was also left out of the description of the apps in our study.

As shown, a multitude of studies applied the UTAUT2 acceptance model in diverse application contexts. However, many researchers also extended and adapted the model to better fit the needs of the specific context of research [18]. The mHealth context is no exception to this [21,24-26]. Different illnesses or ailments have different actual or perceived repercussions. This can result in a stigma for having to deal with the ailment and can cause a fear of losing face if someone were to know about it, and these varying conditions can also affect different needs and perceived necessity in treating the illness, such as taking medication or undergoing physical treatments. Therefore, general acceptance models like the UTAUT2 can only be cautiously applied to the health care context, and extensions of the original predictors need to be considered [11,12].

Additional Constructs

One major barrier to the use of digital and connected technologies is privacy concerns [29]. As mHealth apps also collect and analyze sensitive and intimate personal data, privacy concerns have been shown to be one important impediment to their acceptance [14,30,31].

Privacy can be defined as users' rights to control the flow of personal information [32]. Many users feel that they have lost exactly this control over their personal information in their interaction with digital technologies [33]. They worry about malware, hackers, and identity theft as well as perceived privacy intrusion, secondary use of personal information, and perceived surveillance [34,35]. To understand what shapes privacy concerns and what consequences privacy concerns have, Smith et al [36] proposed the Antecedents–Privacy Concerns–Outcomes macromodel. It shows that privacy concerns are shaped by individual influences (ie, demographic differences and personality differences), experiences, awareness, and culture. Additionally, privacy concerns are also dependent on contextual factors [37].

A large body of research shows that privacy concerns negatively influence users' intention to provide information [36,38]. Correspondingly, privacy concerns represent a barrier to the adoption of technologies that need personal data [29]. However, widely-used technology acceptance models, like the UTAUT2, have not yet integrated privacy concerns, and no new models that integrate privacy concerns have been established. In some empirical studies, privacy concerns were added to the established models, and could improve the prediction of acceptance in different contexts (eg, mobile banking [39], smart city technologies [40], and e-commerce [41]). Also, regarding health information technologies in general and mHealth, privacy concerns have been identified as an important factor to extend established acceptance models in qualitative [12,42,43] and quantitative [16,27,29,44,45] empirical research.

XSL•FO
RenderX

Based on these empirical results, we proposed hypothesis 7: Privacy concerns influence the intention to use mHealth apps.

Another important factor for the acceptance of information technologies and mHealth is trust [14,46,47]. Trust comes into effect in situations of uncertainty and can be described as the attitude to accept this uncertainty and vulnerability based on positive expectations [48,49]. The level of trust is based on the perceived trustworthiness of the technology [50], which is shaped by the reliability and predictability of the technology, the perceived intention of the developers, as well as the individual familiarity with the system [51].

In line with research on privacy concerns and acceptance, it could be shown that technology acceptance models should be extended by trust as an influencing factor, for example, regarding travel apps [52], mobile banking [39,53], and e-commerce [41]. For mHealth, trust has already been identified as an important extension to technology acceptance models [14,27].

This leads to our last hypothesis, hypothesis 8: Trust in mHealth apps influences the intention to use mHealth apps.

Differentiation of Lifestyle and Therapy Apps

The available spectrum of mHealth apps is very broad. Apps related to a healthy lifestyle (eg, fitness, diet, and sleep-monitoring apps)—further on called *lifestyle apps*—are already frequently downloaded, especially by younger people [54,55]. Inspired by the "quantified self" movement and a general public awareness for the responsibility of health-conscious life and behavioral habits [56], many people are interested in using such lifestyle apps. Considering that a healthy lifestyle, regular exercise, a healthy diet, and sufficient sleep can prevent diseases, it is very welcome that mHealth apps trigger lively interest and use among end users. However, to relieve the burden on the health care system, it is necessary that all types of health service apps are offered and used; this includes lifestyle apps as well as apps related to the management of chronic and acute health conditions, the monitoring of important parameters, control of and reminders for medication intake, and other processes supporting the treatment of, and life with, chronic and acute physical or psychological conditions. These are herein called *therapy apps* and are available for different health conditions, ranging from, for example, blood pressure diaries over diabetes trackers to depression and anxiety relief.

Even though the market for lifestyle apps is twice as big as for health care apps and therapeutically orientated apps [57], mHealth apps will support the health care system if they are used by large segments of the population. Specifically regarding therapy apps, positive improvements due to the use of mHealth could be observed, for example, for the cardiac rehabilitation process in older adults [5] or the self-management of diabetes and hypertension [6]. So far, however, the quality of mHealth apps, as well as users' intentions to use such systems, is still questionable [58].

Using apps for a general healthy lifestyle may be influenced by different motives and barriers than using apps for therapy for existing illnesses. Initial empirical evidence shows that acceptance patterns differ between different contexts of digital health technologies [45,59] as well as for different mHealth app types [21]. To better understand user acceptance of mHealth, it is important to disentangle potentially different acceptance patterns.

For these reasons, we pose research question 1: How does the influence of the proposed factors on use intention differ between lifestyle and therapy apps?

User Diversity

People are diverse and so are their evaluation and acceptance of technologies. Besides the highly individual perceptions of, for example, performance, privacy, or influences through peers and habit, user acceptance varies depending on users' characteristics. In general, current users of mHealth apps are rather young, female, and highly educated [54,55], but this also varies depending on the app type. In particular, therapy apps targeted to illnesses that are more prevalent in higher age groups have different target groups. The UTAUT2 incorporates age, gender, and experiences as moderating factors on the relationship between use intention and its antecedents [10]. The effects of sociodemographic characteristics have also been confirmed in other research, for example, in that different age groups attribute varying relevance on acceptance factors [16,60].

However, sociodemographic factors, specifically age, may just be carrier variables for the underlying reasons and user characteristics. The adequate know-how of handling mHealth apps, as well as the fit between needs and target groups, are factors that might have an impact on use intention [61]. Therefore, an important factor for the use and acceptance of mHealth is also familiarity and competence with the use of apps, such as digital health literacy [54,62]. Personal dispositions, such as the individual disposition to value privacy and to trust unknown technologies, may influence mHealth acceptance and explain varying importance of trust and privacy for mHealth acceptance [51,63]. In order to advance an understanding of mHealth app acceptance, the impact of these user diversity factors needs to be further examined. As this is not the focus of this study, we complement our analysis with an exploratory examination of the user diversity factors, which does not suffice for this broad topic but may give first hints on the importance of user diversity for mHealth acceptance.

Therefore, we pose research question 2: How do diverse user characteristics like sociodemographics, experience and literacy with mHealth apps, as well as personal dispositions relate to the acceptance of mHealth apps?

Methods

We used an online questionnaire with a between-subject design to assess the opinions and acceptance by the participants of either lifestyle or therapy apps and to evaluate our hypotheses and research questions.

The Questionnaire

In the introduction section of the questionnaire, a brief orientation on the topic of the study was given. The respondents were also reminded of their rights and informed on how the

XSL•FO
RenderX

collected data would be dealt with. We encouraged them to answer freely, as there were neither "correct" nor "incorrect" answers, and we let them know that we were only interested in their perspective on this timely topic. Respondents were also informed that participation was voluntary and that they were free to quit at any time. Before starting the questionnaire, participants gave consent to collection of their data.

In the next part of the questionnaire—cataloguing the participants' characteristics, attitudes, and experiences—their state of health was first assessed. Besides their subjective health status, questions regarding their experiences with different types of health problems were asked (ie, back and joint pain, headaches and migraine, cardiovascular diseases, allergies and food intolerances, metabolic illnesses, dementia, and a non-option). To measure their experiences with apps, the participants indicated their experience with and use of apps in general and with mHealth apps. For the use of digital health apps in particular, users require skills to search, select, appraise, and apply online health information. Therefore, digital health literacy was assessed in the next part of the questionnaire using the instrument by Van Der Vaart and Drossaert [64], which measures operational skills, navigation skills, information searching, evaluating reliability, determining relevance, adding self-generated content, and protecting privacy (see Table 1 [10,34,51,63,64] for an overview of all constructs). Concerning personality and dispositions, the personal disposition to value privacy [63] as well as the propensity to trust [51], the latter of which was adapted to apps, were assessed. At the end of the questionnaire, sociodemographics (ie, age, gender, and education level) were surveyed.

Table 1. Constructs used in the questionnaire with their respective sources.

Constructs	Subconstructs	Source upon which the construct was based
UTAUT2[a] constructs	• Performance expectancy • Effort expectancy • Social influence • Facilitating conditions • Hedonic motivation • Habit (only answered by users) • Behavioral intention (for users) • Behavioral intention (for nonusers)b	Venkatesh et al [10]
Perceived trust	N/A[c]	Körber [51]
Information privacy concerns	• Perceived surveillance • Perceived intrusion • Secondary use of personal information	Xu et al [34]
Digital health literacy	• Operational skills • Navigation skills • Information searching • Evaluating reliability • Determining relevance • Adding self-generated content • Protecting privacy	Van Der Vaart and Drossaert [64]
Disposition to value privacy	N/A	Xu et al [63]
Propensity to trust (adapted to apps)	N/A	Körber [51]

[a]UTAUT2: unified theory of acceptance and use of technology 2.

[b]This construct was adapted from Venkatesh et al [10].

[c]N/A: not applicable; the construct in this row did not have any subconstructs.

In the main part of the questionnaire, participants were randomly assigned to evaluate either lifestyle apps or therapy apps. The evaluation started with introducing the respective mHealth apps. We assessed performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, habit, and behavioral intention using the items by Venkatesh et al [10] with small adaptations to the context (a detailed overview of the items used, translations, and original items is given in Multimedia Appendix 1).

To assess privacy perceptions, the Mobile Users' Concerns for Information Privacy [34] instrument, which consists of the three subdimensions of perceived surveillance, perceived intrusion, and secondary use of personal information, was used and adapted to lifestyle and therapy apps. To measure trust in the respective mHealth apps, items from the subdimensions of reliability and competence as well as trust in automation by Körber [51] were applied and adapted to the context.

All items were assessed on 6-point symmetric Likert scales ranging from 1 (low agreement) to 6 (strong agreement). Items were randomized to prevent biases. The language of the questionnaire was German, as only German participants were recruited; therefore, items were translated into German. For trust, no validated German translation was available. Therefore, the items were forward-translated by a German native speaker and, to test the comprehensibility and correct translations, two authors translated these again back into English. The results

XSL•FO
RenderX

were compared to the original items, and deviations were settled in a discussion. During the translation process, we also adapted the items to the mHealth context and discussed these adaptations.

To assure a realistic and empathetic evaluation of mHealth apps, the formulation of the items was, on the one hand, individually adapted to the participants already being users of such apps or nonusers. For example, one item for behavioral intention was modified in the following way: "I intend to continue using such a [medical or lifestyle] app" for users versus "I intend to use such a [medical or lifestyle] app in the future" for nonusers. Also, only current and prior users were asked to answer questions concerning habit. Moreover, as therapy apps are less widespread and, therefore, less familiar, the description of therapy apps was illustrated by an example of apps that matched the participants' experienced health problems, as in the questionnaire section about state of health. For those participants who had not experienced any of these health problems before, a general description with several examples was used.

Before distributing the study, we pretested the questionnaire and participants reported back on comprehensibility issues. Only after those issues had been eliminated did we start the data acquisition.

Recruitment of the Sample

Participants were recruited from a university seminar and its attendees' social contacts. Participants accessed the questionnaire via a weblink that was given to them. The comparison between the app types used a between-subject design. Thus, each participant either answered the items regarding therapy apps or lifestyle apps. The app type to be evaluated was assigned randomly. The participants volunteered to take part in the study and were not rewarded for their efforts. Data were collected in May and June of 2019.

The recruitment method was chosen with the aim to reach mHealth users as well as nonusers of therapy and lifestyle apps. Additionally, participants of different age groups were recruited. However, in accordance with the technical requirements of mHealth use, only participants with access to the internet and digital devices were targeted. Today, young people, in particular, use mHealth apps [54]. Therefore, an additional aim for recruitment was to reach those people who have the technological access and know-how to use mHealth apps but still have not adopted this technology, despite possible medical conditions (eg, relatives or acquaintances of seminar students).

Data Analysis

The following sections will detail the analysis methods as well as regulations we applied to our data.

Item Analysis

We checked reliability by using Cronbach α and applied a threshold of a>.70 for all scales not included in the structural model (ie, disposition to value privacy, propensity to trust apps, and digital health literacy). Additionally, as some of the translated German items were not validated previously, we conducted an exploratory factor analysis on the model constructs, confirming the validity of the items (Multimedia Appendix 2).

Structural Model

Our research model was tested using partial least squares (PLS) structural equation modeling (SEM). PLS is a component-based SEM method that is suitable for exploratively testing new models [65], such as this extension of the validated UTAUT2. The analysis process was divided into two parts. First, the measurement quality was checked for reliability and validity. Only when the quality of the model was confirmed, the structural model was analyzed and interpreted.

The software SmartPLS (version 3.3) was used for the SEM modeling [66]. It offers the possibility to conduct multigroup analysis (MGA) to test differences in relationships between constructs and between user groups. We used MGA to test differences between the two app types regarding the strength of the relationships.

As we tailored the questionnaire distinctly, using the targeted app examples, we had to ensure that this did not introduce a systematic error between different apps. We first checked whether the correlations between behavioral intention and the predictor variables differed significantly between the eight app examples used. No such differences were prevalent so that, in the final analysis, no differentiation was made between the participants evaluating therapy apps with different health or ailment foci. For all analyses, a significance level of 5% was set.

Correlation Analysis

To describe how demographics and other user characteristics might be associated with our model variables, we used correlation analysis. To deal with suboptimal normality of our data, we used bias-corrected and accelerated bootstrapping [67].

Exclusion of Participants

Of 951 people who started the questionnaire, 799 completed it (84.0% completion rate). Further, 92 participants with a response time shorter than 50% of the median response time (<16 minutes, 17 seconds) were labeled as speeders and excluded. Finally, 707 participants were included in the analysis.

Data Availability

Access to the anonymized data set can be requested on the Open Science Framework repository [68].

Results

The Sample

The demographic characteristics of the sample are depicted in Table 2 and are differentiated by the type of app the participants evaluated. All in all, the sample included German participants between the ages of 16 and 89 years (mean 36.8, SD 17.5) and 428 women out of 707 (60.5%). The demographic characteristics of the participants were evenly distributed between the two app types. Most participants (n=517, 73.1%) possessed a high education level with a general qualification for university entrance.

XSL•FO
RenderX

Table 2. Demographic characteristics of the sample comparing participants evaluating lifestyle apps and therapy apps (N=707).

Characteristic	Participants evaluating lifestyle apps (n=355)	Participants evaluating therapy apps (n=352)
Age (years), mean (SD)	36.4 (18.1)	37.3 (16.8)
Gender, n (%)		
Women	222 (62.5)	206 (58.5)
Men	133 (37.5)	146 (41.5)
Education level		
No certificate	6 (1.7)	9 (2.6)
Certificate of secondary education	25 (7.0)	25 (7.1)
General certificate of secondary education	59 (16.6)	63 (17.9)
General qualification for university entrance	262 (73.8)	255 (72.4)

Most of the 707 participants used digital technologies: 695 participants (98.3%) owned a smartphone. Only 33 participants (4.7%) did not use apps regularly. Correspondingly, the participants' self-rated app familiarity was quite high (mean 4.32, SD 1.36), as rated on a scale from 1 (very low agreement) to 6 (very high agreement). In contrast, the familiarity with health apps was lower (mean 3.44, SD 2.56). Of the 355 participants evaluating the lifestyle mHealth apps, 110 (31.0%) were current users and 82 (23.1%) had used a lifestyle app before. Only 18 (5.1%) of the 352 participants assigned to the therapy app evaluation group were current users, and 18 (5.1%) had used a therapy app before.

Disclosing information about health status was optional in order to not be too invasive regarding the participants' privacy. Most of the 707 participants reported their health status as "good" (n=305, 43.1%), "very good" (n=216, 30.6%), or "excellent" (n=65, 9.2%). Out of 707 participants, 5 (0.7%) reported their health status as "very bad," 25 (3.5%) reported it as "bad," and 83 (11.7%) reported it as "rather bad." Out of 707 participants, 8 (1.1%) chose not to answer. Out of 707 participants, 26.9% (n=190) lived with a chronic illness, 11.9% (n=84) depended on a medical assistive device, and 31.4% (n=222) needed regular checkups with their physician.

On average, the sample showed a neutral propensity to trust apps in general (mean 3.07, SD 0.79) and a slightly stronger than neutral disposition to value privacy (mean 3.88, SD 1.14). The mean digital health literacy was quite high (mean 4.53, SD 0.75).

The Measurement Model

To assess the quality of the measurement model, the guideline by Hair et al [65] was followed. For the reliability of the model, we confirmed internal consistency reliability (composite reliability >0.708) and considered indicator reliability (outer loading >0.7). The outer loading of one of the items for facilitating conditions on the construct was below 0.4. Dropping it improved reliability. Three other items from the constructs habit, perceived surveillance, and facilitating conditions were closely below the recommended threshold of 0.7, yet they were above 0.6, and were kept in the model as they stemmed from validated models, and dropping them decreased the reliability of the remaining model.

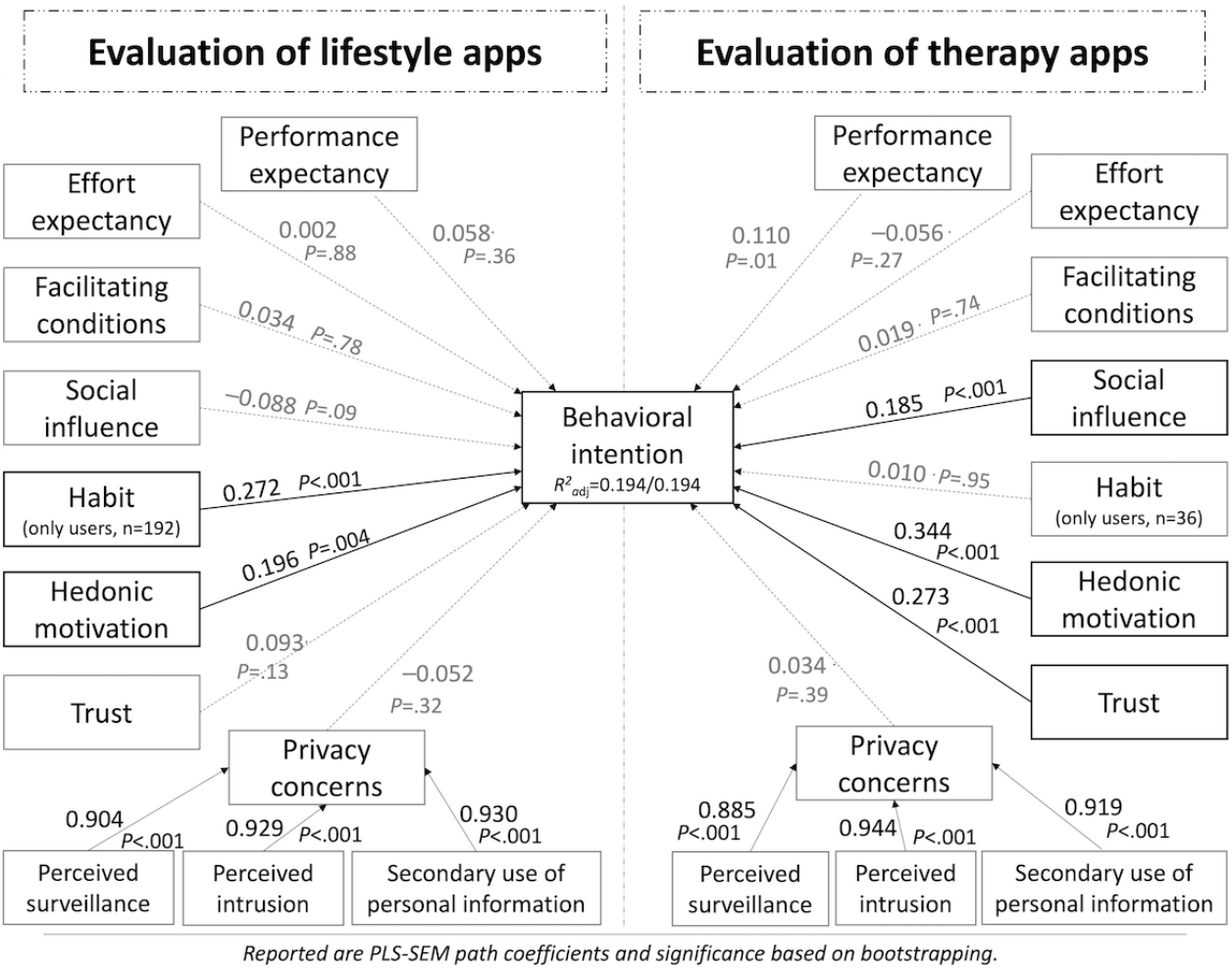
The Structural Model

Evaluation of the validity included convergent validity (average variance extracted >0.5) and discriminant validity, using the Fornell-Larcker criterion. Mobile users' information privacy concerns were modeled as higher-order models because the latent factor privacy concerns was based on three subdimensions. Therefore, validity criteria did not apply to the discriminant validity between the subdimensions themselves or between the subdimensions and the overall scale privacy concerns.

The resulting path coefficients of the model for both lifestyle and therapy apps are depicted in Figure 2. The significance of the path coefficients was checked using bias-corrected and accelerated bootstrapping with 5000 subsamples. Blindfolding procedures were calculated to assess the predictive relevance of the constructs for each app type.

XSL•FO
RenderX

Figure 2. The structural model with path coefficients juxtaposed for lifestyle and therapy apps (significance based on bootstrapping; lifestyle: n=355; therapy: n=352). PLS-SEM: partial least squares structural equation modeling; adj: adjusted.



The results revealed that only 19% of the variance in behavioral intention for both types of mHealth apps could be explained by the extended UTAUT2 model. The variables correspondingly showed only weak predictive relevance for behavioral intentions (Q=0.119 for both app types). Most hypothesized variables showed no significant relationship to behavioral intention. Regarding both types of apps, neither the UTAUT2 constructs of performance expectancy, effort expectancy, and facilitating conditions nor privacy concerns predicted acceptance. Hedonic motivation was the only included construct that had a significant impact on behavioral intention for both app types (lifestyle: 0.196, P=.004, f^2=0.044; therapy: 0.344, P<.001, f^2=0.044). Social influence did impact behavioral intention to use therapy apps (0.185, P<.001, f^2=0.011) but not lifestyle apps. The other way around, habit did impact the intention to use lifestyle apps (0.272, P<.001, f^2=0.02) but not therapy apps. As only current and previous users evaluated habit, the calculation was based on 192 participants for lifestyle apps and 36 participants for therapy apps. In the same vein, trust in the app showed an impact on the intention to use therapy apps (0.273, P<.001, f^2=0.001) but not lifestyle apps.

The MGA confirmed these differences between the evaluation pattern for the app types. Significant differences were present regarding the relationships of habit (D=0.264, P=.002), social influence (D=0.275, P<.001), and trust (D=0.181, P=.04). Table 3 lists the bootstrapped CIs of the path coefficients and the MGA results.

XSL•FO
RenderX

Table 3. Bias-corrected and accelerated bootstrapped 95% CIs for the evaluation of lifestyle and therapy apps and significance of the difference in path coefficients between the two app types based on multigroup analysis (MGA).

Relationship	Lifestyle apps (n=355), 95% CI	Therapy apps (n=352), 95% CI	Significance of MGA, *P* value
Predicting behavioral intention			
Performance expectancy	–0.069 to 0.185	–0.002 to 0.227	.57
Effort expectancy	–0.100 to 0.141	–0.152 to 0.040	.41
Facilitating conditions	–0.125 to 0.123	–0.078 to –0.093	.98
Social influence	–0.197 to 0.013	0.089 to 0.275	<.001
Habit	0.141 to 0.381	–0.126 to 0.106	.002
Hedonic motivation	0.061 to 0.328	0.214 to 0.470	.12
Trust	–0.027 to 0.214	0.146 to 0.399	.04
Privacy concerns	–0.160 to 0.049	–0.041 to 0.110	.19
Higher-order model of privacy concerns			
Perceived surveillance	0.870 to 0.951	0.850 to 0.911	.38
Perceived intrusion	0.885 to 0.951	0.924 to 0.959	.38
Secondary use	0.907 to 0.945	0.894 to 0.939	.51

User Diversity in the Acceptance of mHealth Apps

The validated UTAUT2 model including the additional constructs of privacy concerns and trust showed only weak predictive relevance in explaining why people use mHealth apps and why not. The UTAUT2 postulates that age, gender, and experience moderate the relationships of the predictor variables with behavioral intention [10].

These moderators were not included in our model, as we focused on the direct relationships. Additionally, other human factors have been shown to influence the acceptance of digital technologies and mHealth. Therefore, in an exploratory attempt to decipher how user diversity influences mHealth acceptance and usage, we calculated correlations to get first hints as to what may influence behavioral intention. These results shall not represent a detailed analysis but should give first insights into the impact of selected user characteristics on the acceptance of mHealth apps.

Table 4 depicts the correlations of user factors with behavioral intention to use mHealth apps; these are not differentiated between the two app types. All variables showed significant relationships with behavioral intention. Particularly, familiarity with health apps showed a moderate effect, with a higher familiarity with health apps related to a higher intention for the ongoing use of health apps (r=0.469, P<.001). Additionally, app familiarity (r=0.142, P<.007), propensity to trust apps (r=0.191, P<.001), as well as digital health literacy (r=0.215, P<.001) increased the acceptance of health apps, showing further how important experience and familiarity are to intention for use. Also, demographic characteristics, such as age and gender, showed a significant relationship to acceptance. Older participants and men showed lower acceptance (age: r=–0.15, P<.004; gender: r=–0.075, P=.048), and participants with a higher level of education showed higher acceptance (r=0.195, P<.001). Even though privacy concerns regarding the app itself did not have an impact on behavioral intentions in the structural model, the disposition to value privacy correlated with behavioral intention (r=–0.194, P<.001). Participants who valued their privacy more showed less intention to use mHealth apps.

Table 4. Pearson correlation coefficients of user factors with behavioral intention to use mobile health apps with bias-corrected and accelerated 95% CIs (N=707).

User factor	Correlation with behavioral intention, *r*	95% CI	*P* value
Age	–0.150	–0.255 to –0.034	.004
Gender	–0.075	–0.152 to 0.004	.048
Education level	0.088	0.001 to 0.171	.02
App familiarity	0.142	0.054 to 0.240	.007
Health app familiarity	0.469	0.379 to 0.548	<.001
Digital health literacy	0.215	0.119 to 0.313	<.001
Privacy disposition	–0.194	–0.299 to –0.083	<.001
Propensity to trust apps	0.191	0.88 to 0.291	<.001

XSL•FO
RenderX

Discussion

Overview

The objective of this study was to increase the understanding of users' acceptance and decision patterns to use mHealth apps and which factors impact acceptance for lifestyle apps compared to therapy apps. Therefore, we applied the established technology acceptance model UTAUT2 [10] to the evaluation of mHealth apps. The original model was extended with trust and privacy concerns as predictors of use intention because previous research showed their relevance in the context of digital health technologies (eg, Woldeyohannes and Ngwenyama [27], Lidynia et al [30], and Schomakers et al [59]). In an online questionnaire, 707 participants evaluated either lifestyle or therapy apps in order to compare the decision patterns between these two mHealth app types.

Principal Findings

Overview

In this study, the UTAUT2 model with its extensions can only explain a small amount of variance in the intention to use mHealth apps (approximately 20%). From the original validated UTAUT2 model, only the constructs hedonic motivation, habit, and social influence partly predict the intention to use mHealth apps. The constructs effort expectancy and performance expectancy, which are similarly modeled as main aspects in other established acceptance models like the technology acceptance model [69] and the UTAUT [20], show no significant influence on acceptance. Although researchers already criticized the applicability of these established and widely used acceptance models for the health care context [11,12], our findings deviate from previous empirical research on mHealth app acceptance in which at least some of the factors proposed by the UTAUT2 influenced the intention to use mHealth apps [22,23,25,26]. Therefore, further confirmation of our results is needed, taking into account the limitations of our study, which will be further discussed hereafter, for example, regarding the sample. However, much previous research showed that, on the one hand, not all factors proposed by the UTAUT2 or other technology acceptance models had a significant impact and that, on the other hand, these need to be extended by further influencing factors [22,24,26]. This empirical evidence allows us to conclude that applicability of the UTAUT2 and its predecessors is not fully applicable to the health care and mHealth context, and extended or rather new acceptance models for this context are needed.

Also, in contrast to previous studies (eg, Guo et al [16], Bélanger and Crossler [29], and Schomakers et al [59]), privacy concerns were not found to influence the intention to use mHealth apps. Trust in the reliability and competence of the app, on the other hand, showed a small effect on the acceptance of therapy apps but not on the acceptance of lifestyle apps. Trust and privacy concerns have been extensively studied in information systems research [70-72]. However, for both concepts, no commonly agreed-upon definition and operationalization exist in research. After all, privacy and trust are not completely disjunct (eg, trust beliefs can mitigate privacy concerns) [59,73]. For both reasons, it is important to study different aspects of privacy and trust. Our results suggest that privacy concerns regarding the perceived surveillance, intrusion, and secondary use of information do not impact mHealth acceptance, but maybe concerns regarding hacker attacks, misuse of information by health insurance companies, or similar concerns do. In the same vein, trust in the reliability and competence of mHealth apps showed only a weak influence on the acceptance of therapy apps in our study. Trusting the mHealth app provider, the data protection mechanisms, or a physician recommending the use of an app may have an impact. These other dimensions of trust and privacy need to be further examined while paying close attention to the specific operationalization of the constructs. Therefore, our results are only a first step toward studying the impact of privacy and trust on mHealth acceptance.

Our results further suggest that instead of the more "utilitarian" aspects of perceived usefulness and performance of mHealth apps, it is rather the "emotional" aspects, such as fun, prior experiences, and recommendations by peers, that are important for their use. This finding must be confirmed and further analyzed in future studies, but it indicates that, on the one hand, approaches that address user experience, such as gamification, are important for mHealth apps of both types as the hedonic motivation influenced use intention for both app types. On the other hand, personal and peer experiences are very influential, whereby a widespread use of mHealth apps becomes even more important.

Context Differences

Besides the general model, our results revealed differences in the importance of some predictors for lifestyle and therapy apps. In our sample and in general, lifestyle apps were far more frequently used than therapy apps. The categorization of mHealth apps into lifestyle and therapy apps is not disjunct, as some apps may have functions providing both. In our study, the introduction given to the participants clearly distinguished between apps used to improve fitness, nutrition, and similar for "lifestyle" and those apps providing support for dealing with a prevailing illness. However, for future research, a classification of mHealth apps that is commonly agreed upon is vital, as is the simplification of research on context differences.

Habit emerged as a significant acceptance factor only for the lifestyle apps, which may be explained by the more widespread use and larger proportion of users in the sample. However, as the sample of users for therapy apps was very small (ie, only 36 participants), these results have to be interpreted with caution. The behavioral intention to use therapy apps was, in contrast to lifestyle apps, also influenced by social influence and trust. In this medical context, the participants need more than fun to use the app and, rather, should search for more reliable and trustworthy apps. Similar results have been found by Schomakers et al [21].

User Diversity

All in all, the predictive relevance of the factors in the extended UTAUT2 model is rather weak. This confirms other authors' opinions regarding health care technologies, in that the established models can only be cautiously applied and need

XSL•FO
RenderX

further adaptations, or rather, new models for the special health care context are needed (eg, Ziefle and Wilkowska [9] and Venkatesh [10]).

The results of our exploratory analysis of the relationship between different user factors and behavioral intention to use mHealth apps imply that user diversity is an important aspect that needs to be considered. In particular, experience showed a strong relationship with use intention in this preliminary analysis, particularly the experience with health apps, but also with apps in general. The same was true for digital health literacy. Further empirical research and analysis of user diversity, especially the importance of experience, is needed, but this first result hints at a developing acceptance. When more and more people use mHealth apps, including therapy apps, the increased familiarity combined with habit and social influence may increase acceptance within the population. On the other hand, it cannot be assumed that everybody has the experience and the skills to use mHealth apps. Digital health literacy has to be developed and, following the gray digital divide, older people in particular, who are also more prone to chronic conditions, need support in getting to know these digital helpers.

Limitations

Despite the valuable insights into decision patterns regarding mHealth apps, this online questionnaire approach needs to be considered methodologically. Instead of actual adoption behavior, *reported* attitudes, perceptions, and intention to adopt were measured for mHealth apps in general, not regarding a specific mHealth app. Therefore, more general implications can be extracted from the results; however, on the other hand, users evaluated a vague idea of what mHealth apps are. Furthermore, their evaluations may be strongly influenced by those mHealth apps they have already experienced, especially as experience was shown to be strongly related to acceptance. This adds variability to the data. Therefore, this general research should be accompanied by more research into app-specific acceptance, which can provide detailed results for the optimal design of apps. Moreover, as an urgent research desideratum, the role of prior experience needs to be further explored. While the difference between users and nonusers [74] of mHealth apps might impact technology acceptance, so too could the usage of different mHealth apps probably influence future acceptance of mHealth apps in general. Additionally, the rather young and educated sample needs to be considered, which was acquired via social contacts. Convenience sampling has the advantage that those people actually participating are often highly motivated to provide their opinion. However, by their motivation and their self-selection, bias may have been brought into the data.

Besides the young and healthy persons still improving their health via lifestyle apps, thereby preventing chronic diseases, very important target groups for mHealth are older people and people with health problems. Here, mHealth can unfold its potential in directly supporting therapy and monitoring diseases, thereby improving quality of care and relieving the health care systems in a short time. These user groups should be further researched as they are underrepresented within our sample. Also, the German nationality of the participants limits the implications from this research, as attitudes toward technologies are highly influenced by cultures (eg, Trepte et al [75] and Alagöz et al [76]).

As no validated translation of the UTAUT2 items to German was known to us when planning the study, and no validated adaptation to the health care context was yet available, the use of unvalidated translations and adaptations of the items might have lowered the validity of our results. We could statistically assure a good validity and reliability of our items; nevertheless, the use of validated scales is highly recommended for future research (for a German translation of UTAUT2 items, see Harborth and Pape [77]; for a validated French adaptation to eHealth technologies, see Hayotte et al [78]).

Conclusions

In this study, an extended UTAUT2 technology acceptance model was used to predict behavioral intention to use mHealth apps. Only a few hypothesized predictors (ie, hedonic motivation, habit, and social influence) showed a significant relationship to use intention, and the model only explained a comparably small amount of variance (approximately 20%). These factors indicate that more emotional factors than utilitarian usefulness influence mHealth app acceptance, adding a piece to the understanding of the mHealth acceptance puzzle. Small differences in the decision patterns were prevalent between the acceptance of lifestyle apps (eg, for fitness, nutrition, and sleep) and therapy apps (eg, for the monitoring and treatment of back pain, migraine, and cardiovascular diseases). In future research, the results need to be replicated, as the generalizability from our rather young sample is limited. However, our results in combination with previous research indicate that the UTAUT2 model, which was developed for the acceptance and use of mobile internet technologies in general, is not very suitable to predict mHealth use. The health care context needs improved and adapted technology acceptance models, which must also include human factors, such as experience, to account for user diversity.

Acknowledgments

We would like to thank all participants for openly sharing their opinions.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Questionnaire items, sources, and translations.
[DOCX File , 23 KB - mhealth_v10i1e27095_app1.docx]

Multimedia Appendix 2
Exploratory factor analysis.
[DOCX File , 21 KB - mhealth_v10i1e27095_app2.docx]

References

1. Lin TTC, Bautista JR. Understanding the relationships between mHealth apps' characteristics, trialability, and mHealth literacy. J Health Commun 2017 Apr;22(4):346-354. [doi: 10.1080/10810730.2017.1296508] [Medline: 28323546]
2. Morris ME, Aguilera A. Mobile, social, and wearable computing and the evolution of psychological practice. Prof Psychol Res Pr 2012 Dec;43(6):622-626 [FREE Full text] [doi: 10.1037/a0029041] [Medline: 25587207]
3. van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: A research agenda for the field of m-health. Bull World Health Organ 2012 May 01;90(5):393-394 [FREE Full text] [doi: 10.2471/BLT.11.099788] [Medline: 22589575]
4. Messner EM, Probst T, O'Rourke T, Stoyanov S, Baumeister H. mHealth applications: Potentials, limitations, current quality and future. In: Baumeister H, Montag C, editors. Digital Phenotyping and Mobile Sensing. Cham, Switzerland: Springer; 2019:235-248.
5. Bostrom J, Sweeney G, Whiteson J, Dodson JA. Mobile health and cardiac rehabilitation in older adults. Clin Cardiol 2020 Feb;43(2):118-126 [FREE Full text] [doi: 10.1002/clc.23306] [Medline: 31825132]
6. Choi W, Wang S, Lee Y, Oh H, Zheng Z. A systematic review of mobile health technologies to support self-management of concurrent diabetes and hypertension. J Am Med Inform Assoc 2020 Jun 01;27(6):939-945 [FREE Full text] [doi: 10.1093/jamia/ocaa029] [Medline: 32357368]
7. Rossmann C, Krömer N. mHealth in der medizinischen versorgung, prävention und gesundheitsförderung. In: Fischer F, Krämer A, editors. eHealth in Deutschland. Berlin, Germany: Springer Vieweg; 2016:441-456.
8. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: A systematic review and meta-analysis. Ann Intern Med 2015 Jan 20;162(2):123-132. [doi: 10.7326/M14-1651] [Medline: 25599350]
9. Ziefle M, Wilkowska W. Technology acceptability for medical assistance. In: Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare.: IEEE; 2010 Presented at: 4th International ICST Conference on Pervasive Computing Technologies for Healthcare; March 22-25, 2010; Munich, Germany. [doi: 10.4108/ICST.PERVASIVEHEALTH2010.8859]
10. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178. [doi: 10.2307/41410412]
11. Holden RJ, Karsh B. The technology acceptance model: Its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
12. Vassli LT, Farshchian BA. Acceptance of health-related ICT among elderly people living in the community: A systematic review of qualitative evidence. Int J Hum Comput Interact 2017 May 30;34(2):99-116. [doi: 10.1080/10447318.2017.1328024]
13. Bansal G, Zahedi F, Gefen D. Do context and personality matter? Trust and privacy concerns in disclosing private information online. Inf Manage 2016 Jan;53(1):1-21. [doi: 10.1016/j.im.2015.08.001]
14. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What predicts patients' adoption intention toward mHealth services in China: Empirical study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]
15. Risch S. Gesundheit in Zahlen [Health in Numbers]. Schwerpunkt: Risiko. Nr. 2. Germany: IKK classic; 2018 Oct. URL: https://www.ikk-classic.de/assets/885_ikkc_web_pdf.pdf [accessed 2022-01-06]
16. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. Electron Commer Res Appl 2016 Mar;16:55-65. [doi: 10.1016/j.elerap.2015.11.001]
17. Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley Publishing Company, Inc; 1975.
18. Tamilmani K, Rana N, Wamba S, Dwivedi R. The extended unified theory of acceptance and use of technology (UTAUT2): A systematic literature review and theory evaluation. Int J Inf Manage 2021 Apr;57:102269. [doi: 10.1016/j.ijinfomgt.2020.102269]
19. Castanha J, Pillai SKB, Indrawati. What influences consumer behavior toward information and communication technology applications: A systematic literature review of UTAUT2 model. In: Proceedings of the 5th International Conference on Information and Communication Technology for Sustainable Development. 2021 Presented at: 5th International Conference on Information and Communication Technology for Sustainable Development; July 23-24, 2020; Goa, India p. 317-327. [doi: 10.1007/978-981-15-8289-9_30]

XSL•FO
RenderX

20. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: Toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
21. Schomakers E, Lidynia C, Ziefle M. Exploring the acceptance of mHealth applications - Do acceptance patterns vary depending on context? In: Proceedings of the 9th International Conference on Applied Human Factors and Ergonomics.: Springer; 2019 Presented at: 9th International Conference on Applied Human Factors and Ergonomics; July 22-26, 2018; Orlando, FL p. 53-64. [doi: 10.1007/978-3-319-94619-1_6]
22. Sun Y, Wang N, Guo X, Peng Z. Understanding the acceptance of mobile health services: A comparison and integration of alternative models. J Electron Commer Res 2013;14(2):183-200 [FREE Full text]
23. Yuan S, Ma W, Kanthawala S, Peng W. Keep using my health apps: Discover users' perception of health and fitness apps with the UTAUT2 model. Telemed J E Health 2015 Sep;21(9):735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
24. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
25. Wang H, Liang L, Du C, Wu Y. Implementation of online hospitals and factors influencing the adoption of mobile medical services in China: Cross-sectional survey study. JMIR Mhealth Uhealth 2021 Feb 05;9(2):e25960 [FREE Full text] [doi: 10.2196/25960] [Medline: 33444155]
26. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: Survey study and quantitative model development. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e17588 [FREE Full text] [doi: 10.2196/17588] [Medline: 32673249]
27. Woldeyohannes H, Ngwenyama O. Factors influencing acceptance and continued use of mHealth apps. In: Proceedings of the 4th International Conference on HCI in Business, Government and Organizations.: Springer; 2017 Presented at: 4th International Conference on HCI in Business, Government and Organizations; July 9-14, 2017; Vancouver, BC p. 239-256. [doi: 10.1007/978-3-319-58481-2_19]
28. Sánchez-Fernández R, Iniesta-Bonillo M. The concept of perceived value: A systematic review of the research. Mark Theory 2007 Dec;7(4):427-451. [doi: 10.1177/1470593107083165]
29. Bélanger F, Crossler RE. Privacy in the digital age: A review of information privacy research in information systems. MIS Q 2011;35(4):1017-1041. [doi: 10.2307/41409971]
30. Lidynia C, Brauner P, Ziefle M. A step in the right direction – Understanding privacy concerns and perceived sensitivity of fitness trackers. In: Proceedings of the 8th International Conference on Applied Human Factors and Ergonomics.: Springer; 2018 Presented at: 8th International Conference on Applied Human Factors and Ergonomics; July 17-21, 2017; Los Angeles, CA p. 42-53. [doi: 10.1007/978-3-319-60639-2_5]
31. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: Meta-ethnographic review of qualitative studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/13817] [Medline: 31293246]
32. Altman I. Privacy: A conceptual analysis. Environ Behav 2016 Jul 26;8(1):7-29. [doi: 10.1177/001391657600800102]
33. European Commission. Special Eurobarometer 431: Data Protection. Brussels, Belgium: Directorate-General for Communication; 2015. URL: https://data.europa.eu/data/datasets/s2075_83_1_431_eng?locale=en [accessed 2022-01-06]
34. Xu H, Gupta S, Rosson M, Carroll J. Measuring mobile users' concerns for information privacy. In: Proceedings of the 33rd International Conference on Information Systems. 2012 Presented at: 33rd International Conference on Information Systems; December 16-19, 2012; Orlando, FL URL: https://faculty.ist.psu.edu/xu/papers/Xu_etal_ICIS_2012a.pdf
35. Paine C, Reips U, Stieger S, Joinson A, Buchanan T. Internet users' perceptions of 'privacy concerns' and 'privacy actions'. Int J Hum Comput Stud 2007 Jun;65(6):526-536. [doi: 10.1016/j.ijhcs.2006.12.001]
36. Smith H, Dinev T, Xu H. Information privacy research: An interdisciplinary review. MIS Q 2011;35(4):989-1015. [doi: 10.2307/41409970]
37. Nissenbaum H. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA: Stanford University Press; 2010.
38. Li Y. Empirical studies on online information privacy concerns: Literature review and an integrative framework. Commun Assoc Inf Syst 2011;28:453. [doi: 10.17705/1cais.02828]
39. Merhi M, Hone K, Tarhini A. A cross-cultural study of the intention to use mobile banking between Lebanese and British consumers: Extending UTAUT2 with security, privacy and trust. Technol Soc 2019 Nov;59:101151. [doi: 10.1016/j.techsoc.2019.101151]
40. Habib A, Alsmadi D, Prybutok VR. Factors that determine residents' acceptance of smart city technologies. Behav Inf Technol 2019 Nov 21;39(6):610-623. [doi: 10.1080/0144929x.2019.1693629]
41. Sánchez-Torres JA, Arroyo-Cañada FJ, Rojas-Berrio SP, Robayo-Pinzon OJ, Fontalvo-Cerpa W. The Colombian electronic consumer: Analysis of the leading factors of e-commerce use. Int J Electron Mark Retailing 2019;10(3):283. [doi: 10.1504/IJEMR.2019.100704]
42. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: A systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
43. Spann A, Stewart E. Barriers and facilitators of older people's mHealth usage: A qualitative review of older people's views. Hum Technol 2018 Nov 30:264-296. [doi: 10.17011/ht/urn.201811224834]

XSL•FO
RenderX

44. Halbach P, Himmel S, Offermann-van Heek J, Ziefle M. A change is gonna come. The effect of user factors on the acceptance of ambient assisted living. In: Proceedings of the 4th International Conference, ITAP 2018: Human Aspects of IT for the Aged Population, Acceptance, Communication and Participation.: Springer; 2018 Presented at: 4th International Conference, ITAP 2018: Human Aspects of IT for the Aged Population, Acceptance, Communication and Participation; July 15-20, 2018; Las Vegas, NV p. 52. [doi: 10.1007/978-3-319-39943-0]
45. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Ind Manage Data Syst 2015;115(9):1704-1723. [doi: 10.1108/IMDS-03-2015-0087]
46. Meng F, Guo X, Peng Z, Lai K, Zhao X. Investigating the adoption of mobile health services by elderly users: Trust transfer model and survey study. JMIR Mhealth Uhealth 2019 Jan 08;7(1):e12269 [FREE Full text] [doi: 10.2196/12269] [Medline: 30622092]
47. Ghazizadeh M, Peng Y, Lee JD, Boyle LN. Augmenting the technology acceptance model with trust: Commercial drivers' attitudes towards monitoring and feedback. In: Proceedings of the 56th Human Factors and Ergonomics Society Annual Meeting. 2012 Presented at: 56th Human Factors and Ergonomics Society Annual Meeting; October 22-26, 2012; Boston, MA p. 2286-2290. [doi: 10.1177/1071181312561481]
48. Rousseau D, Sitkin S, Burt R, Camerer C. Not so different after all: A cross-discipline view of trust. Acad Manage Rev 1998 Jul;23(3):393-404. [doi: 10.5465/amr.1998.926617]
49. Lee JD, See KA. Trust in automation: Designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
50. Akter S, D'Ambra J, Ray P. Trustworthiness in mHealth information services: An assessment of a hierarchical model with mediating and moderating effects using partial least squares (PLS). J Am Soc Inf Sci Technol 2010 Oct 25;62(1):100-116. [doi: 10.1002/asi.21442]
51. Körber M. Theoretical considerations and development of a questionnaire to measure trust in automation. In: Proceedings of the 20th Congress of the International Ergonomics Association.: Springer; 2018 Presented at: 20th Congress of the International Ergonomics Association; August 26-30, 2018; Florence, Italy p. 13-30. [doi: 10.1007/978-3-319-96074-6_2]
52. Gupta A, Dogra N, George B. What determines tourist adoption of smartphone apps? J Hosp Tour Technol 2018 Mar 12;9(1):50-64. [doi: 10.1108/jhtt-02-2017-0013]
53. Alalwan AA, Dwivedi YK, Rana NP. Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. Int J Inf Manage 2017 Jun;37(3):99-110. [doi: 10.1016/j.ijinfomgt.2017.01.002]
54. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: A source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
55. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
56. Lupton D. Quantifying the body: Monitoring and measuring health in the age of mHealth technologies. Crit Public Health 2013 Dec;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
57. Wang A, An N, Lu X, Chen H, Li C, Levkoff S. A classification scheme for analyzing mobile apps used to prevent and manage disease in late life. JMIR Mhealth Uhealth 2014 Feb 17;2(1):e6 [FREE Full text] [doi: 10.2196/mhealth.2877] [Medline: 25098687]
58. Powell AC, Landman AB, Bates DW. In search of a few good apps. JAMA 2014 May 14;311(18):1851-1852. [doi: 10.1001/jama.2014.2564] [Medline: 24664278]
59. Schomakers E, Lidynia C, Ziefle M. Listen to my heart? How privacy concerns shape users' acceptance of e-Health technologies. In: Proceedings of the 15th International Conference on Wireless and Mobile Computing, Networking and Communications.: IEEE; 2019 Presented at: 15th International Conference on Wireless and Mobile Computing, Networking and Communications; October 21-23, 2019; Barcelona, Spain p. 306-311. [doi: 10.1109/wimob.2019.8923448]
60. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. Int J Inf Manage 2018 Dec;43:342-350. [doi: 10.1016/j.ijinfomgt.2017.08.006]
61. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: Experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
62. Smith B, Magnani JW. New technologies, new disparities: The intersection of electronic health and digital health literacy. Int J Cardiol 2019 Oct 01;292:280-282 [FREE Full text] [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]
63. Xu H, Dinev T, Smith HJ, Hart P. Examining the formation of individual's privacy concerns: Toward an integrative view. In: Proceedings of the 29th International Conference on Information Systems. 2008 Presented at: 29th International Conference on Information Systems; December 14-17, 2008; Paris, France.
64. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: Measuring a broad spectrum of Health 1.0 and Health 2.0 skills. J Med Internet Res 2017 Jan 24;19(1):e27 [FREE Full text] [doi: 10.2196/jmir.6709] [Medline: 28119275]
65. Hair Jr JF, Hult GTM, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). 2nd edition. Thousand Oaks, CA: SAGE Publications; 2017.

XSL•FO
RenderX

66. Ringle C, Wende S, Becker JM. SmartPLS. Bönningstedt, Germany: SmartPLS GmbH; 2015. URL: https://www.smartpls.com/ [accessed 2019-01-01]
67. Field A. Discovering Statistics Using IBM SPSS Statistics. 4th edition. Thousand Oaks, CA: SAGE Publications; 2015.
68. User acceptance of mHealth apps - Applying an extended UTAUT2 model to explain acceptance of lifestyle and therapy apps. Open Science Framework. 2021 Mar 22. URL: https://osf.io/rkm7n/?view_only=d097fe38650e49ca80d213388021f7cc [accessed 2022-01-07]
69. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
70. Abu-Seman S, Ramayah T. mHealth application security and privacy concerns: A comprehensive review of past literature. Adv Sci Lett 2017 Sep 01;23(9):8957-8960. [doi: 10.1166/asl.2017.10003]
71. Alan AK. The role of trust in mobile technologies usage in emerging countries. In: Mtenzi FJ, Oreku GS, Lupiana DM, Yonazi JJ, editors. Mobile Technologies and Socio-Economic Development in Emerging Nations. Hershey, PA: IGI Global; 2018:234-261.
72. Sarkar S, Chauhan S, Khare A. A meta-analysis of antecedents and consequences of trust in mobile commerce. Int J Inf Manage 2020 Feb;50:286-301. [doi: 10.1016/j.ijinfomgt.2019.08.008]
73. Rohm A, Milne G. Just what the doctor ordered. J Bus Res 2004 Sep;57(9):1000-1011. [doi: 10.1016/s0148-2963(02)00345-4]
74. Sezgin E, Özkan-Yildirim S, Yildirim S. Intention vs perception: Understanding the differences in physicians' attitudes toward mobile health applications. In: Sezgin E, Yildirim S, Özjkan-Yildirim S, Sumuer E, editors. Current and Emerging mHealth Technologies. Cham, Switzerland: Springer; 2018:153-166. [doi: 10.1007/978-3-319-73135-3_10]
75. Trepte S, Reinecke L, Ellison N, Quiring O, Yao M, Ziegele M. A cross-cultural perspective on the privacy calculus. Soc Media Soc 2017 Jan 01;3(1):1-13 [FREE Full text] [doi: 10.1177/2056305116688035]
76. Alagöz F, Ziefle M, Wilkowska W, Valdez A. Openness to accept medical technology - A cultural view. In: Proceedings of the Symposium of the Austrian HCI and Usability Engineering Group. 2011 Presented at: Symposium of the Austrian HCI and Usability Engineering Group; November 25-26, 2011; Graz, Austria p. 151-170. [doi: 10.1007/978-3-642-25364-5_14]
77. Harborth D, Pape S. German translation of the unified theory of acceptance and use of technology 2 (UTAUT2) questionnaire. SSRN J 2018 Mar 28:1-12 [FREE Full text] [doi: 10.2139/ssrn.3147708]
78. Hayotte M, Thérouanne P, Gray L, Corrion K, d'Arripe-Longueville F. The French eHealth Acceptability Scale using the unified theory of acceptance and use of technology 2 model: Instrument validation study. J Med Internet Res 2020 Apr 15;22(4):e16520 [FREE Full text] [doi: 10.2196/16520] [Medline: 32293569]

Abbreviations

MGA: multigroup analysis
mHealth: mobile health
PLS: partial least squares
SEM: structural equation modeling
UTAUT2: unified theory of acceptance and use of technology 2

Edited by L Buis; submitted 11.01.21; peer-reviewed by E Sezgin, H Meggy, S Liu, R Bond, R De Croon; comments to author 03.03.21; revised version received 09.04.21; accepted 24.09.21; published 18.01.22.

Please cite as:
Schomakers EM, Lidynia C, Vervier LS, Calero Valdez A, Ziefle M
Applying an Extended UTAUT2 Model to Explain User Acceptance of Lifestyle and Therapy Mobile Health Apps: Survey Study
JMIR Mhealth Uhealth 2022;10(1):e27095
URL: https://mhealth.jmir.org/2022/1/e27095
doi:10.2196/27095
PMID:35040801

©Eva-Maria Schomakers, Chantal Lidynia, Luisa Sophie Vervier, André Calero Valdez, Martina Ziefle. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 18.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing Elderly User Preference for Telehealth Solutions in China: Exploratory Quantitative Study

Nuoya Chen[1], PhD; Pengqi Liu[2], MSc

[1]Faculty of Global Studies, Justice and Rights, University of Macerata, Macerata, Italy
[2]Sino-Danish College, University of Chinese Academy of Sciences, Beijing, China

Corresponding Author:
Nuoya Chen, PhD
Faculty of Global Studies, Justice and Rights
University of Macerata
Crescimbeni 30-32
Macerata, 62100
Italy
Phone: 39 0733 2582418
Email: cny0824@gmail.com

Abstract

Background: In the next 15 to 20 years, the Chinese population will reach a plateau and start to decline. With the changing family structure and rushed urbanization policies, there will be greater demand for high-quality medical resources at urban centers and home-based elderly care driven by telehealth solutions. This paper describes an exploratory study regarding elderly users' preference for telehealth solutions in the next 5 to 10 years in 4 cities, Shenzhen, Hangzhou, Wuhan, and Yichang.

Objective: The goal is to analyze why users choose telehealth solutions over traditional health solutions based on a questionnaire study involving 4 age groups (50-60, 61-70, 71-80, and 80+) in 4 cities (Shenzhen, Hangzhou, Wuhan, and Yichang) in the next 10 to 20 years. The legal retirement age for female workers in China is 50 to 55 years and 60 years for male workers. To simulate reality in terms of elderly care in China, the authors use the Chinese definition of elderly for employees, defined as being 50 to 60 years old rather than 65 years, as defined by the World Health Organization.

Methods: The questionnaires were collected from Shenzhen, Hangzhou, Wuhan, and Yichang randomly with 390 valid data samples. The questionnaire consists of 31 questions distributed offline on tablet devices by local investigators. Subsequently, Stata 16.0 and SPSS 24.0 were used to analyze the data. O-logit ordered regression and principal component analysis (PCA) were the main theoretical models used. The study is currently in the exploratory stage and therefore does not seek generalization of the results.

Results: Approximately 71.09% (280/390) of the respondents reported having at least 1 type of chronic disease. We started with PCA and categorized all Likert scale variables into 3 factors. The influence of demographic variables on Factors 1, 2, and 3 was verified using analysis of variance (ANOVA) and *t* tests. The ordered logit regression results suggest that health-related motivations are positively related to the willingness to use telehealth solutions, and trust on data collected from telehealth solutions is negatively correlated with the willingness to use telehealth solutions.

Conclusions: The findings suggest that there is a need to address the gap in community health care and ensure health care continuity between different levels of health care institutions in China by providing telehealth solutions. Meanwhile, telehealth solution providers must focus on improving users' health awareness and lower health risk for chronic diseases by addressing lifestyle changes such as regular exercise and social activity. The interoperability between the electronic health record system and telehealth solutions remains a hurdle for telehealth solutions to add value in health care. The hurdle is that doctors neither adjust health care plans nor diagnose based on data collected by telehealth solutions.

(JMIR Mhealth Uhealth 2022;10(1):e27272) doi:10.2196/27272

KEYWORDS

telehealth solutions; preference; motivation; elderly user; China

XSL•FO
RenderX

Introduction

Background

COVID-19 has severe effects on the elderly population with multiple chronic diseases such as hypertension, diabetes, and cardiovascular diseases than the healthy subpopulation. Mortality rate analysis shows higher death rates caused by COVID-19 associated with those aged above 50 years [1]. Population projections suggest that the Chinese population will peak from 2025 to 2030 [2], thereby leading to surging demands for high-quality medical services including telehealth solutions. Research on the preference of the elderly for telehealth solutions must be conducted. The legal retirement ages for female and male employees in China are 55 and 60 years respectively, ranking as one of the lowest in the world [3]. Therefore, to study elderly user preferences, the authors chose to start from those aged 50 years to simulate reality at the best possible level in this study.

Telehealth refers to the use of telecommunication tools for health care continuum in the prevention, treatment, diagnosis, recovery, and home care processes. The use of wearables and apps for health management and online hospitals for health consultation has become increasingly popular with the wide use of smartphones. COVID-19 has accelerated the digitalization of the health care system at a pace unimaginable a few years ago. The use of telemedicine services has increased by more than 1000% in March and more than 4000% in April [4]. Spending on the use of telehealth solutions also increased starting from March by more than 1000 % [5].

There is a need to explore user preference for telehealth solutions, particularly for the future generations of elderly (aged 50 years and above) who will become 70 years old by 2030. The Chinese society is facing challenges posed by a rapidly aging population and rising chronic disease trends caused by lifestyle changes owing to the urbanization process. Based on projections, the Chinese population will peak between 2026 and 2030 [2]. Thus, researching whether and how telehealth solutions can generate more value for the elderly in the next 15 to 20 years is important. Owing to the lack of high-quality medical resources and trained clinicians, there is an urgent need to look for alternative solutions such as telehealth solutions. The implementation of telehealth solutions faces challenges among elderly users given their lack of experience with technology and thereby the lack of trust in telehealth solutions. Other factors such as the household income, education level, and health status of the users may also play a role.

The rest of the paper is structured as follows. The next section provides the literature review on research methodologies used to study users' willingness for using telehealth solutions. We then present the methodology and research design used for the analysis, followed by the description of the qualitative and quantitative analyses of the questionnaire revealing why users choose telehealth solutions over traditional health solutions. Finally, we summarize the main findings and implications of this study.

Literature Review

To analyze the state of the art of the research methodology regarding user and physician preferences for telehealth solutions, a thorough literature review was conducted.

There have been several empirical studies on patients in all age groups and clinician perceptions regarding telehealth solutions. The multinominal logit regression model has become popular for statistical analyses in health economics and marketing science [6]. The paired *t* test has also been also used for comparing the preference for traditional health visits with telehealth consultations or the presence of telephysicians.

Direct-to-consumer telehealth solutions roughly comprise 3 categories [6]. The first category covers solutions provided by the same doctor from whom the patients obtain primary care services. As the health care service is provided by the doctor with whom patients have established a relationship, telehealth solutions can ensure convenience to patients while maintaining care continuity. The second category incorporates solutions provided by doctors from the same institution where patients receive health care services but not the same doctor with whom the patients have an established relationship. This allows the patients' records to be updated by the doctor from the same care institution while maintaining the connection with the care home. Meanwhile, patients can receive care during and after working hours. The third category consists of telehealth solutions provided by doctors who have no previous relationship with the patients or the patients' primary care service providers. Many newly emerging telehealth solution providers belong to the third category. Patients can pay for the services provided out of their pockets or by claiming insurance.

One study [7] used SurveyMonkey to send out a questionnaire for conducting a nation-wide survey in the United States. In total, there were 4345 patients covering different ethnicities, age groups, and income groups with various education backgrounds and insurance coverages. The survey aimed to determine the willingness of the participants to use telehealth solutions and their comfort level with telehealth solutions belonging to the aforementioned 3 categories of solutions. Results from the generalized estimation equation model showed that patients were more willing to use category I solutions. The willingness to use telehealth solutions declined if the provider had no relationship with patients before or if the services were provided by other doctors from the same care institutions. More than half of the patients were willing or very willing to use telehealth solutions involving their own doctor. One-third of all participants were willing to use telehealth solutions provided by other doctors from the same care institution. Less than 20% of all participants were willing to use telehealth solutions provided by doctors with whom they had no previous relationship. Patients' comfort in using telehealth solutions grows with the attachment to their original care institution.

In another study [8], the authors have tried to analyze patient preferences and satisfaction rates with the telehealth program, CVS Minute Clinics. Minute Clinics offer patients video consultation with doctors at collaboration clinics while assisting nurses in performing on-site diagnostic tests and using tools such as otoscopes, telephonic stethoscopes, and digital video

XSL•FO
RenderX

laryngoscopes to assist doctors in making diagnoses by reading the image or data on the screen. Such treatment costs US $59 on average with life insurance [9].

The survey participants were over 18 years old and agreed to use telemedicine service when on-site doctors were busy. The study used the logistic regression model to assess the preferences of 1734 users of the Minute Clinics services. Among these participants, 94% to 99% reported high satisfaction with telehealth solutions. One-third of all participants preferred telehealth solutions to traditional health solutions. The authors suggest that the lack of medical insurance, gender (female) of the users, self-satisfaction with the understanding of telehealth solutions, service quality, and convenience can predict user preference for telehealth visits [10]. Patients' satisfaction with on-site nurses has an adverse relationship with the preference for telehealth solutions. The possible explanations are that the more satisfied patients are with on-site nurses, the more they are reminded of the benefits of in-person interactions. Moreover, patients may get the false impression that on-site nurses alone can perform the necessary diagnosis and therefore ignore the fact that on-site nurses do not have the license to practice alone.

One study [11] has analyzed factors associated with clinicians' perceptions regarding telehealth solutions and examined if these factors affected their decision to continue using telehealth solutions after COVID-19. Doctors from different disciplines, including pediatricians and doctors focusing on adult patients, surgical and nonsurgical doctors, outpatient and inpatient doctors, and doctors who focus on both categories have been covered [10]. The 220 full responses also covered doctors with and without previous telehealth experiences. The study disseminated a Likert scale questionnaire and used logistic regression to analyze the odds of different factors at a significance level of 95%.

Results [10] suggest that ease of use for patients is the most important feature followed by ease of use for clinicians. Physicians' overall satisfaction [11] and perceived ease of use [12] also directly affect perceived usefulness and the intention to use telemedicine. Meanwhile, the quality of care, ease of physical examination, and beliefs on whether adaptability is an important quality of clinicians also play a role in determining doctors' preference for telehealth solutions. Being more perceiving rather than judging is also seen as one of the personality factors affecting clinicians' decision to extend their use of telehealth solutions. Moreover, clinicians' beliefs regarding the importance of physical touch have a negative correlation with their decision to extend the use of telehealth solutions.

The study conducted by Miner [10] suggests that clinicians play a significant role in adapting to the digital health trends. Training may prove necessary to help clinicians continue their telehealth practices after COVID-19.

Researchers [13] have studied outpatients' use of the internet to search for orthopedic information. They used a questionnaire consisting of 12 questions that was distributed by doctors to outpatients during office visits. A total of 1161 complete responses were collected and analyzed with a multivariable binominal logistic regression model. Regression results show that younger age groups are primarily associated with increased use of the internet for obtaining health and orthopedic information. Younger patients are also more likely to find the search results related to their current orthopedic problems "very helpful" and "somewhat helpful." Google is a more popular search engine than Yahoo and Bing. Patients who visited sports medicine clinics were less likely to use WebMD to search for answers to their orthopedic questions. Other than this, the type of clinic did not have a significant effect on patients' use of the internet. Males were more likely to find information from the internet very useful than female patients; besides this, gender does not have a significant impact on patients' internet usage. The study suggests that patients seem to conduct research on the internet with search engines more than on the website of the institution where they are being treated.

Another study [14] confirms that using the internet for searching information along with telehealth solutions, and doctors' suggestions in clinics and hospitals shall address the problem where patients rely on search engines to search answers to medical problems because of the lack of reliable medical information sources online. Chatbots can offer an alternative for such a problem.

The study [14] compared the accuracy of traditional nurse triages and physician telepresence at an emergency pediatric department. The study used paired *t* tests to analyze the triage time and accuracy (triage utility) differences between traditional nurse triages and physician telepresence. In total, data on 100 families were collected in this study, which took place at a large, tertiary care children's hospital with 65,000 emergency department visits occurring annually. Physician telepresence was achieved using the RP-7i robot, with a built-in stethoscope after the patients went through the traditional nurse triage. The questionnaire consists of 9 5-point Likert scale questions and 1 yes/no question to assess the overall experience of using the robot.

At P=.10, there is no difference in the triage time between the traditional nurse triage and physician telepresence. There are statistically significant differences between the triage accuracy of traditional nurse triages and physician telepresence (P=.03). The triage accuracy score of the traditional nurse triage is at 71% whereas the physician telepresence score is at 95%. Parents and children have preference scores for physician telepresence and indicate that they would choose physician telepresence during their next pediatric emergency department visit [14].

Another study [15] focused on children who were 5.99 years old on average. These children preferred new technology. In the emergency department, time is everything whereas it may be tricky for nurses to make accurate judgments without enough physicians in the emergency room. The robotic experience has significantly improved triage accuracy by avoiding missing values on the triage form, which consists of 27 items. This suggests that in an overwhelmed emergency room, having physician telepresence may help ease stress and avoid mistakes.

The study [15] also analyzed the impact of the integrated health care buddy project with patients having chronic disease conditions in the United States. The study is a collaboration study between 2 clinics at Washington and Oregon, Robert

XSL•FO
RenderX

Bosch Healthcare and American with 2 groups of patients (an intervention group and a control group), each comprising 1767 patients with chronic obstructive pulmonary disease, congestive heart failure, or diabetes. The health buddy program gives a free handheld device for patients to use at home and a large screen. The device connects patients with care managers and allows patients to interact with their care managers about vital signs, symptoms, health-related knowledge, and behavior. Insurance claim data were used to analyze the cost for managing chronic disease and mortality rates.

Moreover, the study confirms the effectiveness of harnessing telehealth assistants for chronically ill patients. Telehealth solutions not only lowered the mortality rates by 2.7% in the intervention group over 2 years but also saved costs between 7.7% and 13.3% per patient per quarter (US $312-542). The study used multivariate regression to predict the cost reduction for patients who engaged more with the program and patients who do not engage otherwise. The prediction showed cost savings of US $1009 per congestive heart failure patient per quarter (P<.001). For patients engaged in the program, the cost saving is US $968 per patient per quarter (P<.001). For patients who did not engage with the program, the cost saving is not significant. For hospital admission rates, the study suggests that telehealth intervention lowers inpatient admission by 3.4% (P<.001).

The paper highlights the need to recognize the value of integrated telehealth solutions for high-risk patients with chronic diseases who incur high costs. Having a device at home for allowing patients to interact with care managers not only allows care managers to capture the deteriorating vital signs and provide interventions in time but also to identify gaps in patients' health knowledge and behavior [15].

The discrete choice experiment (DCE) has been a popular tool to identify the preference for telehealth solutions and the different attributes related to the preference for such solutions [16]. Researchers tried to identify the preference of elderly people (aged 65 years and above) in Australia. The study [16] analyzed factors such as the distance to the nearest clinic and cost of virtual visits and their influence on the preference level for telehealth solutions. The study indicates that most of the elderly have never used the internet in the past 3 months, indicating a knowledge gap for elderly users in using telehealth solutions. In the study, 330 respondents were recruited with a mean age of 69 years. The study concludes that participants would rather use telehealth solutions only as complementary tools with in-person visits. As the study was conducted in Adelaide, Australia, where the age structure, family structure, and health status of the elderly are different from those in China, there is a need to analyze the preference of elderly users for telehealth in China.

In another study [17], the choice between mobile health and telehealth was studied with the DCE model involving 1403 residents in rural areas. The study suggests that the preference is associated with the gender and setting of the users. The distance (access to health care) to hospitals and their gender determines if the residents would prefer using telehealth solutions.

Methods

Questionnaire Distribution

With the legal retirement age standing at 55 years for female employees and 60 years for male employees in China, the questionnaire was distributed among the future elderly (aged above 50 years) in Shenzhen, Hangzhou, Wuhan, and Yichang to best simulate reality. This study followed the DCE methodology and comprised 5 stages including designing research questions, interviews with experts, interviews with individual users, the pretest stage, and the pilot test stage. The questionnaire study was conducted with assistance from the University of Chinese Academy of Sciences and Beijing Cinsos Consulting Corporation. We collected 390 valid answers from 50-60, 60-70, 70-80, and 80+ age groups to analyze individual users' willingness to use telehealth solutions over traditional health solutions.

Ethical Approval

Ethical approval was obtained in May 2019 from the committee in University of Macerata.

Based on the ethical approval results and the analytical results from focus group analysis, questionnaires were designed to analyze stakeholders' attitudes in China toward whether the Internet of Healthcare Things solutions can help reduce the gap in the demands of the current health care system.

Table 1 summarizes the collected data. All data were stored on the "Box" owned by KU Leuven.

Table 1. Data collection summary (source: author's design).

Data	Format	Transfer	Consent	Pseudonymization
Recording with consumers	Windows Media Audio (WMA), MP3, MP4	Data were collected for scientific research purposes and therefore transferred from China to Europe and stored on cloud.	Question 1 in the questionnaire (see Multimedia Appendix 1)	Yes
Questionnaire collected on tablet devices	Word	Data were collected for scientific research purposes and therefore transferred from China to Europe and stored on cloud.	Question 1 in the questionnaire (see Multimedia Appendix 1)	Yes
Excel form with summary of data pseudonymized	Excel	Data were collected for scientific research purposes and therefore transferred from China to Europe and stored on cloud.	Question 1 in the questionnaire (see Multimedia Appendix 1)	Yes

XSL•FO
RenderX

Study Design

The main purpose of the survey was to understand the factors affecting the preference of elderly users for telehealth solutions. The DCE model based on the random utility theory to evaluate the preference for telehealth solutions [18] was used. The DCE method is widely used in studying how patients value different attributes of health care services and the potential demand for new services or treatment [19]. The study follows the standard DCE methodology, namely (1) defining research questions to compile evidence, (2) interviewing experts (stakeholders), (3) interviewing individual users(focus group studies), (4) the pretest stage (online questionnaire in Europe, N=31), and (5) the pilot test stage (online questionnaire in Xiangyang, China, N=104). In the pilot test stage, 104 questionnaires were answered, with 55 questionnaires containing usable data (mostly from Hubei Province).

The questionnaire (see Multimedia Appendix 1) consists of 31 questions and 5 parts. The questionnaire starts with a screening question on whether the participant is willing to participate in the survey and share data for scientific research purposes. There are 10 Likert scale questions related to the motivation, 7 questions surrounding the demographic information including participants' insurance coverage, 6 questions about the usage of telehealth solutions at the time of survey, 4 questions about the health status of survey participants, and 3 questions about whether users want to share data with insurance companies, doctors from community health centers, and doctors from hospitals. The degree of influence of each factor is evaluated with a Likert scale from 1 to 7 (1=no influence, 4=neutral, and 7=with influence). The questionnaire was written in Chinese and then translated in English for easy understanding by the author.

The questionnaire has 5 parts; the first part is about the current situation of telehealth solution usage by the surveyed elderly users.

Telehealth solutions are defined as smartphone apps (such as Alihealth, Ping An Good Doctor, Chun Yu Doctor, Wedoctor, Yue Dong Quan, etc), wearables (such as Xiaomi Band, Huawei watch and Apple Watch, etc), health management tools for home use (such as PICOOC smart scale, Mi Home i-Health blood pressure monitor, Mi Home Hi-Pee Smart Pee Monitor, Smart Sleep Monitor, Smart devices to improve sleep quality, etc.). The section consists of 4 questions asking the usage frequency, reasons for starting to use telehealth solutions, if telehealth solutions were used to monitor sleep, and if telehealth solutions were used to monitor nutrition.

The second part of the questionnaire is about the health status of the survey participants (self-evaluated). The third part asks about the potential benefits of telehealth solutions and elderly users' motivations. The fourth part is designed around the potential risks of telehealth solutions (price, privacy risk, data accuracy risk, brand and design, resistance to technology, and usage experience). The fifth part is designed to gather demographic information, including gender, age, residence, household income, and education in years.

Further, 13 questions were designed focusing on the reasons for preferring telehealth solutions to traditional health solutions. The following questions are related to F2, the perceived benefits of telehealth solutions: monitoring health status (Q13), reducing health risks (Q14), following the doctor's advice (Q15), free devices provided by insurance companies (Q18), and lack of community health care services (Q20).

There were also questions regarding the perceived risk for telehealth solutions (F3), including data accuracy (trust) concerns (Q22), privacy concerns (Q23), financial reasons for the price (Q24), design, popularity, and usage difficulty concerns (Q25).

As some of the reasons for using telehealth solutions pertain to the social image of the individuals, social influence (Q28) is also considered one of the factors that could influence users' preference.

Data Collection

In China, the legal retirement age for female factory workers is 50 years, 55 years for female employees, and 60 years for male employees. To simulate reality with respect to elderly care in China, elderly is defined as being over 50 years old instead of being 65 years old according to the World Health Organization. In our study, we intended to compare the participants' willingness to use telehealth solutions considering different age groups and residents in different cities, with the data collection target set for each age group (50-60, 61-70, and 71-80 years) containing approximately 100 data subjects. Data subjects more than 80 years old were categorized as being in the 50-100 years group because of the health conditions that limited the number of participants.

In the pretest stage of the study, questionnaires in English were distributed on the internet via Microsoft Forms through the Philips intranet portal and Berlin Expat Group on Facebook. We collected 31 questionnaires. In the pilot testing stage, the questionnaire was translated into Chinese and distributed via the internet with Wenjuanxing through WeChat. We collected 104 questionnaires with 55 valid answers. The pretesting stage was designed to test the design of the questionnaire; therefore, the data collected were not analyzed.

In the distribution stage, questionnaires were disseminated on tablet devices by local investigators randomly among residents more than 50 years old in Shenzhen, Hangzhou, Wuhan, and Yichang with the help of Beijing Cinso Consulting Corporation. More than 450 questionnaires were distributed, and 402 answers were collected, with a recovery rate of 89%. Among them, 390 were completely valid questionnaires, accounting for 87% of the questionnaires issued and 97% of all the questionnaires returned. The other 12 questionnaires were not used in data analysis because they did not provide complete information or were deemed to have not been filled carefully.

The data were collected in Chinese language and then summarized in an Excel sheet (Microsoft Corporation) and converted into a pseudonymized value form in Excel. Data were then analyzed using SPSS 24.0 (IBM Corporation) and Stata 16.0 (StataCorp).

XSL•FO
RenderX

The level of urban development differs with Tier 1, 2, 3, and 4 cities; the disposable income of residents in the designated cities and the medical resources accessible (hospitals and doctors) vary as well (see Table 2). This may lead to differences in the preference for telehealth solutions.

Table 2. Disposable income in Shenzhen, Hangzhou, Wuhan, and Yichang (source: CEIC, 2020; National bureau of statistics, 2020).

Category	City	GDP[a] in 2019 (billion US$)	Disposable and discretionary income (US$)
Tier 1	Shenzhen	422.875	9818.83
Tier 2	Hangzhou	241.425	10357.65
Tier 3	Wuhan	254.744	8120.17
Tier 4	Yichang	70.05	4518.5

[a]GDP: gross domestic product.

Shenzhen was chosen because it is the headquarter of Ping An Technology. Ping An Technology has worked with the government of Shenzhen and other stakeholders to provide electronic medical insurance schemes. Residents in Shenzhen can now use the Ping An Good Doctor app to buy complementary insurance in addition to the basic medical insurance schemes and get refunded online.

Hangzhou was chosen, as it is the city where Alibaba is headquartered. During the interview with Alihealth, the fact that 80% of all primary health care facilities in Zhejiang Province are now equipped with artificial intelligence–assisted image recognition systems was mentioned.

Wuhan was chosen, as it is an important hub in Central China where the population is growing rapidly in recent years. Recently, the Wuhan Municipality has launched several programs promoting the internet + home care initiative for the elderly. There are several exploratory projects running in different districts in Wuhan such as in the Dongxihu and Wuchang districts. There have been several models proposed and tested in Wuhan for elderly care such as the community embedded model, centralization model, and combinations of the proper centralization and decentralization models. Services provided to the elderly focus on assisted food service, assisted cleaning service, assisted nursing and medical service, and long-distance care.

Yichang was chosen, as the level of aging population in Yi Chang is higher than the national average. Aging was measured by the percentage of people over 60 years old in the entire population and the percentage of people over 80 years old in the elderly population. The Yichang municipality is currently developing community-based care centers and rural cooperative elderly care centers. The Yichang municipality established the telehealth solution platform for elderly care in 2019. Using the platform, in 2020, the tele-elderly-care (translated from Chinese) services were expected to reach all townships in Yichang and cover over 50% of all elderly people.

Theoretical Model and Hypothesis

To evaluate users' willingness to use telehealth solutions, 3 hypotheses were formulated; in addition, the model considers the effects of demographic factors such as age, education background, income, health status, and living habits such as regular social activity and regular exercise. Table 3 describes the theoretical model built to assess users' willingness to use telehealth solutions, the hypothesis, and the variables involved in the model. The theoretical model consists of 2 parts; the first part assesses the Likert scale factors and their correlation with the users' willingness; the second part assesses demographic factors and their impact on the 3 factors and the willingness to use telehealth solutions.

Considering that the dependent variable, namely the willingness to use telehealth solutions, is an ordered discrete variable, the ordered logit model is used for regression. The impact of each factor was assessed by designing 4 models.

$$Y = \beta F_1 + \gamma Z + \varepsilon \quad \textbf{(1)}$$

$$Y = \beta F>_2 + \gamma Z + \varepsilon \quad \textbf{(2)}$$

$$Y = \beta F_3 + \gamma Z + \varepsilon \quad \textbf{(3)}$$

$$Y = \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \gamma Z + \varepsilon \quad \textbf{(4)}$$

Model (1) is used to test the impact of Factor 1, and models (2) and (3) are used to test the impact of Factors 2 and 3. Model (4) considers the influence of the above 3 factors.

Y represents the designated value for the willingness of participants to use telehealth solutions. In the original questionnaire, the question assigned the preference level as from 1 to 7 (1=preference for traditional health solutions [face-to-face communication], 4=neutral, and 7=preference for telehealth solutions). *Z* represents the control variables such as demographic factors, including the living city, age, gender, education level, health condition, income, living situation, and lifestyle variables (regular exercise and regular social activity) of the participants.

XSL•FO
RenderX

Table 3. Hypotheses and corresponding variables in the model.

Factor	Hypothesis	Corresponding question in the questionnaire
Factor 1		
1.1 Social Influence	1.1 Social influence (friend and family opinions) has an impact on the willingness to use telehealth solutions.	Q28, Likert scale value: 1-7
1.2 Price	1.2 The price of telehealth solutions has an impact on the willingness to use telehealth solutions.	Q24, Likert scale value: 1-7
1.3 Design and brand	1.3 The brand and design of telehealth solutions have an impact on the willingness to use telehealth solutions.	Q25, Likert scale value: 1-7
1.4 Privacy risk	1.4 The privacy risk associated with the use of telehealth solutions has an impact on the willingness to use these solutions.	Q23, Likert scale value: 1-7
1.5 Private insurance or business insurance coverage	1.5 Private or business insurance plan coverage has an impact on the willingness to use telehealth solutions.	Q18, Likert scale value: 1-7
Factor 2: Health-related motivation factors		
2.1 Lower health risk	2.1 The belief that telehealth solutions can lower health risk is positively related to the willingness to use telehealth solutions.	Q14, Likert scale value: 1-7
2.2 Raise health awareness	2.2 The belief that telehealth solutions can raise health awareness is positively related to the willingness to use telehealth solutions.	Q13, Likert scale value: 1-7
2.3 Lack of community health care for patients	2.3 The belief that telehealth solutions can amend the gap in the lack of community health care for patients has an impact on the willingness to use telehealth solutions.	Q22, Likert scale value: 1-7
2.4 Unstable doctor-patient relationship	2.4 The belief that telehealth solutions can help improve doctor-patient relationship has an impact on the willingness to use telehealth solutions.	Q15, Likert scale value: 1-7
Factor 3: Trust		
3. Data accuracy	3. The accuracy of the data collected by telehealth solutions has an impact on the willingness to use telehealth solutions.	Q22, Likert scale value: 1-7
Control variables		
Residence city	The residence city of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q3, 1=Shenzhen, 2=Hangzhou, 3=Wuhan, and 4=Yichang
Gender	The gender of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q30, 0=female and 1=male
Education	The education level of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q31, 1=Primary school education (0-6 years), 2=junior/senior high school education (6-12 years), 3=vocational training (12-15 years), 4=college education (15-18 years), and 5=graduate school education (>=18 years)
Income	The monthly household income of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q29, 1=no fixed income, 2=monthly household income <=US $785.23, 3=monthly household income >US $785.23 but <=US $1570.45, 4=monthly income >US $1570.45) but <=US $4711.35, and 5=monthly income >US $4711.35, original value in RMB, 1 USD=6.37 RMB
Health status	The self-reported health status of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q11, 1=self-reported healthy, 2=suboptimal healthy, 3=with chronic disease having no significant impact on life quality, and 4=have chronic disease with significant impact on life quality

XSL•FO
RenderX

Factor	Hypothesis	Corresponding question in the questionnaire
Preferred living status	The preferred living situation of the participants has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q4, 1=Prefer living alone, 2=prefer living with partner, 3=prefer living with children, and 4=prefer living with children or grandchildren
Regular exercise	Participants engaging or not engaging in regular exercise has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q9, 1=Socialize regularly and –1=do not socialize regularly
Regular social activity	Participants engaging or not engaging in regular social activities has an impact on Factors 1, 2, and 3 and their willingness to use telehealth solutions.	Q10, 1=Exercise regularly and –1=do not exercise regularly

Results

Health Status of Survey Participants

Based on the self-identified responses from the subjects, the following categories were created to identify their health status: healthy, suboptimal healthy, with chronic disease, and self-identified healthy. Then, more detailed data, such as the type of chronic disease and the number of chronic diseases of the survey participants, were analyzed.

Among the 390 participants, 117 (30%) reported having 1 chronic disease (30%), 64 (16.4%) reported having 2 chronic diseases (16.4%), and 47 (12.05%) responded as having 3 chronic diseases (12.05%). Further, 17 participants (4.36%) stated having 4 chronic diseases (4.36%), 7 (1.79%) reported having 5 chronic diseases, 6 (1.53%) reported having 6 chronic diseases, and 2 (0.5%) reported having communicable and chronic diseases (0.5%). Furthermore, 110 participants (28%) reported having no chronic diseases.

Descriptive Statistics

In this section, the qualitative analytical results are presented. All the survey participants are over 50 years old because the survey intends to collect information on elderly users' needs in the next 5 to 10 years, as shown in Table 4. Among the 390 participants providing valid answers, 160 (41.03%) indicate that they are more willing to use traditional health care solutions; 167 (42.82%) indicate that they are willing to use telehealth solutions, whereas 51 (13.07%) show neutral willingness.

Among the 390 participants, 112 (28.7%) are aged 51 to 60 years; another 112 participants (28.7%) are aged 61 to 70 years. Further, 110 participants (28.2%) are aged 71 to 80 years is. There are 43 participants (11.4%) over 80 years old; the number of participants in this group is less than that in the other 3 age groups, as data subject recruitment was limited by the physical conditions of the individuals in this age group.

Moreover, 67.7% (264/390) of the participants often use telehealth solutions to monitor health status. Most survey participants (246/390, 63.1%) received 6 to 12 years of education, followed by 83 participants (21.3%) who went to elementary school. Given the survey candidate recruitment conditions for the elderly aged above 50 years, the education level of the participants is in line with the reality.

The distribution diagram in Figure 1 shows that participants in the age group of 51 to 60 years and those aged above 80 years show a strong willingness or a willingness to use telehealth solutions. This may be because people in the age group of 50 to 60 years are more familiar with technology, whereas those above 80 years cannot physically attend in-person doctor visits.

XSL•FO
RenderX

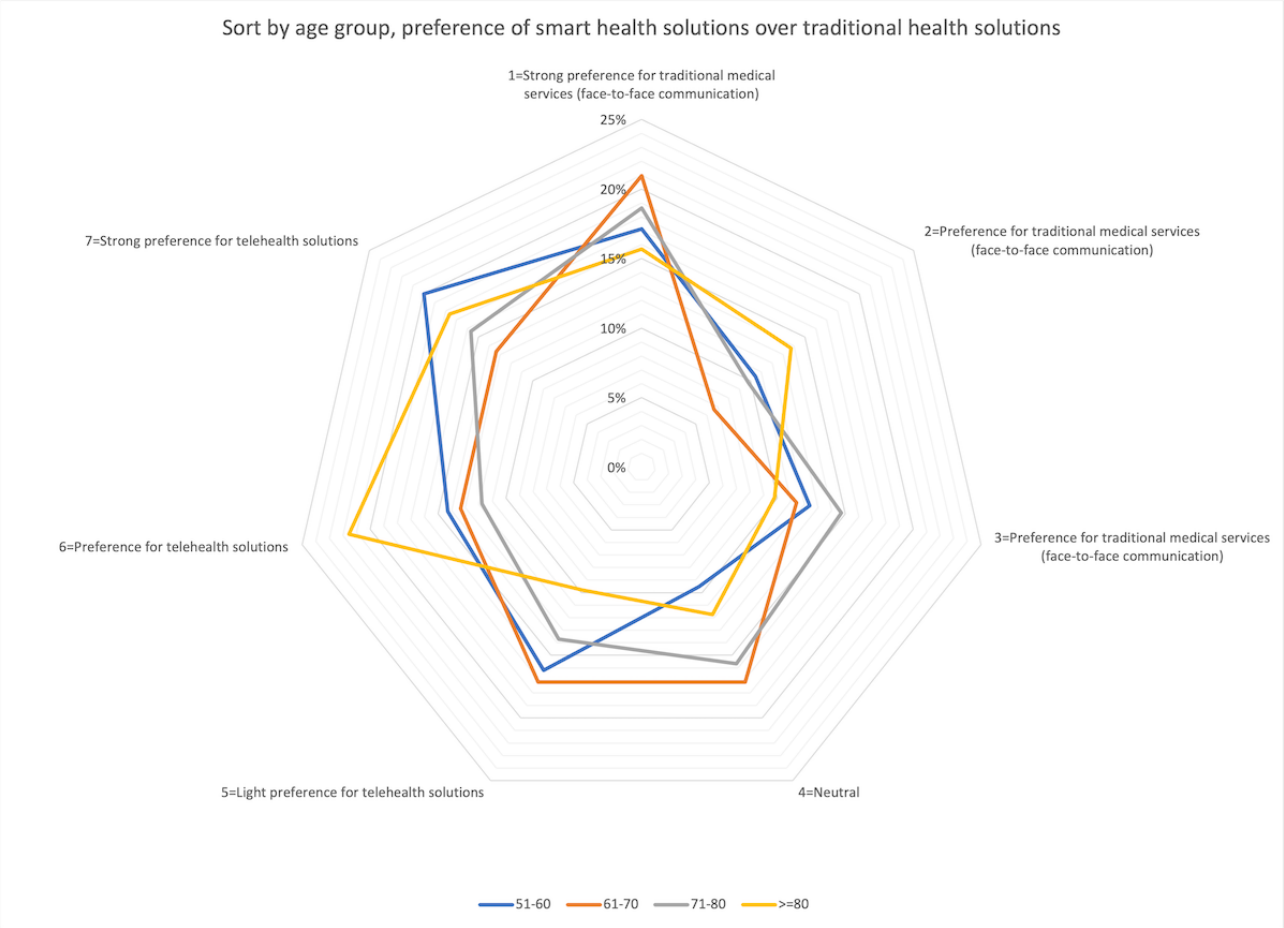
Table 4. Demographic characteristics of participants (N=390).

Characteristic	n (%)
Gender	
Male	224 (57.4)
Female	166 (42.6)
Age (years)	
51-60	112 (28.7)
61-70	112 (28.7)
71-80	110 (28.2)
>=80	56 (14.4)
Residence city	
Shenzhen	97 (24.9)
Hangzhou	95 (24.4)
Wuhan	108 (27.7)
Yichang	90 (23.1)
Household income (US$, original value in RMB, $1 = 6.37 RMB)	
No fixed monthly income	21 (5.4)
≤785.23	84 (21.5)
785.23-1570.45	186 (47.7)
1570.45-4711.35	88 (22.6)
≥4711.35	11 (2.8)
Frequency of using telehealth solutions	
Often	264 (67.7)
Occasionally	82 (21)
Rarely	44 (11.3)
Education level	
Primary school (1-6 years)	83 (21.3)
Junior or high school (6-12 years)	246 (63.1)
Vocational training (12-15 years)	31 (7.9)
College graduate (15-18 years)	29 (7.4)
Graduate School (>=18 years)	1 (0.3)
Health status	
Healthy	145 (37.2)
Suboptimal healthy	99 (25.4)
With minor chronic disease	132 (33.8)
With major chronic disease affecting life quality	14 (3.6)
Living situation	
Living alone	47 (12.1)
Living with partner	156 (40)
Living with children	177 (45.4)
Living with grandchildren	4 (1)
Health insurance status	
None	8 (2.1)
Basic resident or employee medical insurance	309 (79.2)

XSL•FO
RenderX

Characteristic	n (%)
Private insurance	7 (1.8)
Other social insurance schemes	21 (5.4)
Public servant insurance	38 (9.7)
Unknown	7 (1.8)

Figure 1. Willingness to use telehealth solutions sorted by age group.



Among the 390 users surveyed, there are 224 males and 166 females, accounting for 57.4% and 42.6% of the total number of participants, respectively; the proportion of male users is higher than that of female users (as observed in Table 4). Figure 2 suggests that female users are willing to use traditional medical solutions, whereas male users are strongly willing to use telehealth solutions.

In accordance with the study design, survey participants are evenly distributed in the 4 cities. The number of data samples obtained in Shenzhen, Hangzhou, Wuhan, and Yichang are 97, 95, 108, and 90, accounting for 24.9%, 24.4%, 27.7%, and 23.1%, respectively, of the 390 data subjects. Figure 3 suggests that in Shenzhen and Wuhan, the percentage of participants showing preference for telehealth solutions is higher than that in Hangzhou and Yichang.

The distribution of income follows the bell curve, with approximately half (186/390, 47.7%) of the sample's monthly household income falling between US$ 785.23 and US$ 1570.45; the proportions of the sample with household monthly incomes less than or equal to US$ 785.23 and more than or equal to US$ 4711.35 account for only 5.4% (21/390) and 2.8% (11/390), respectively. The willingness to use telehealth solutions increases with the monthly income as well. Figure 4 points out that in the >=US$ 4711.35 income group, the preference is mainly neutral and above neutral. The lower the income, the higher the percentage of the surveyed data subjects showing strong preference for traditional health solutions. This can be observed among the no income and <=US$ 785.23 income groups.

In terms of using telehealth solutions for monitoring sleep and nutrition intake, the percentage of users who are currently using telehealth solutions for sleep monitoring and nutrition monitoring are respectively 23.07% (90/390) and 26.15% (102/390).

Figure 5 indicates that the major reasons for using telehealth devices are self-care, following the doctor's advice, and the free devices and services offered by insurance companies in China or a mix of these 3 reasons.

XSL•FO
RenderX

Figure 2. Willingness to use telehealth solutions based on gender.

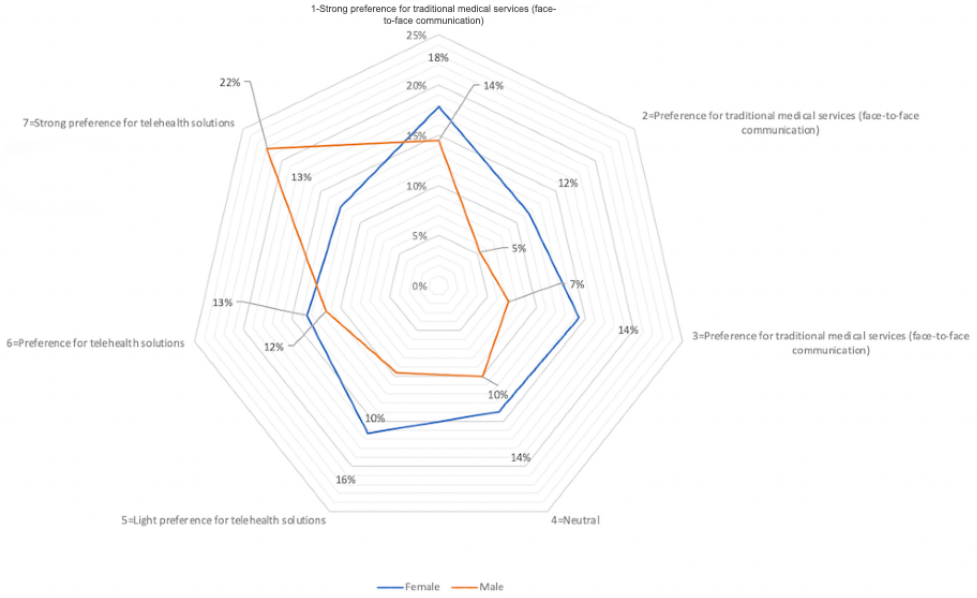
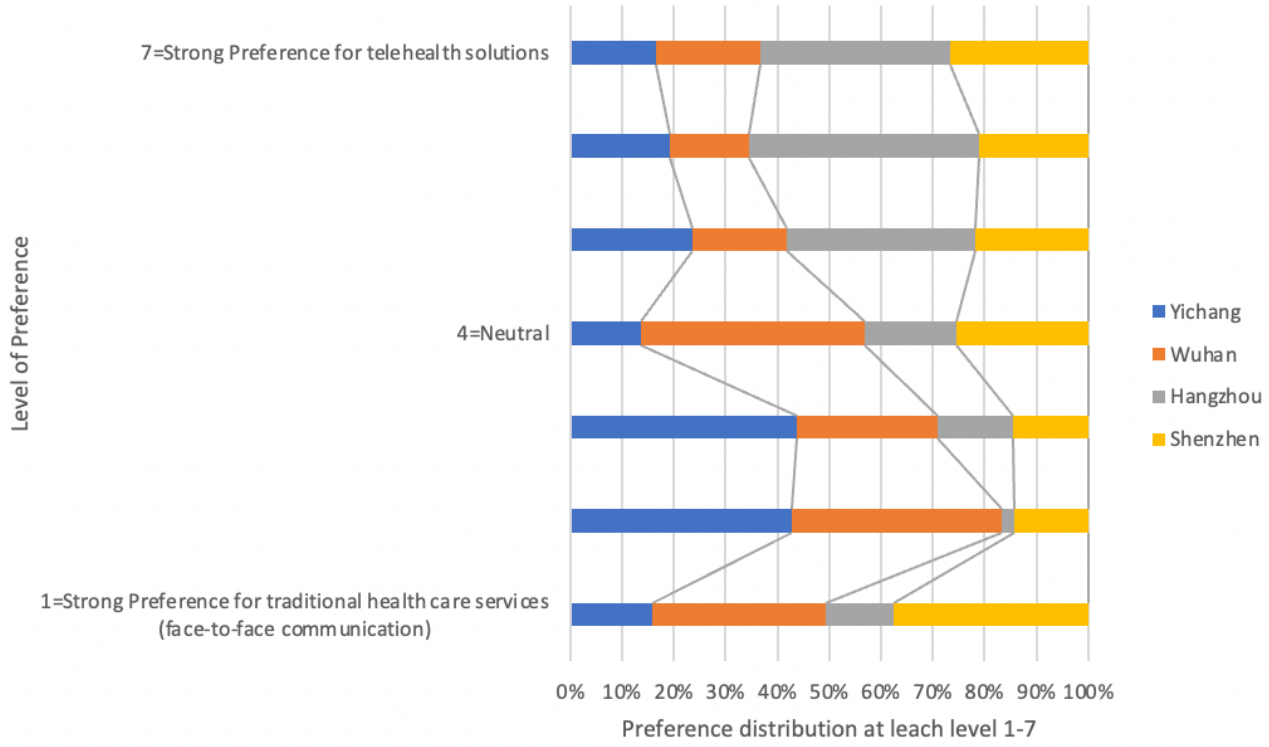


Figure 3. Willingness to use telehealth solutions in Yichang, Wuhan, Hangzhou, and Shenzhen.



XSL•FO
RenderX

Figure 4. Willingness of the participants in the 5 income groups to use telehealth solutions.

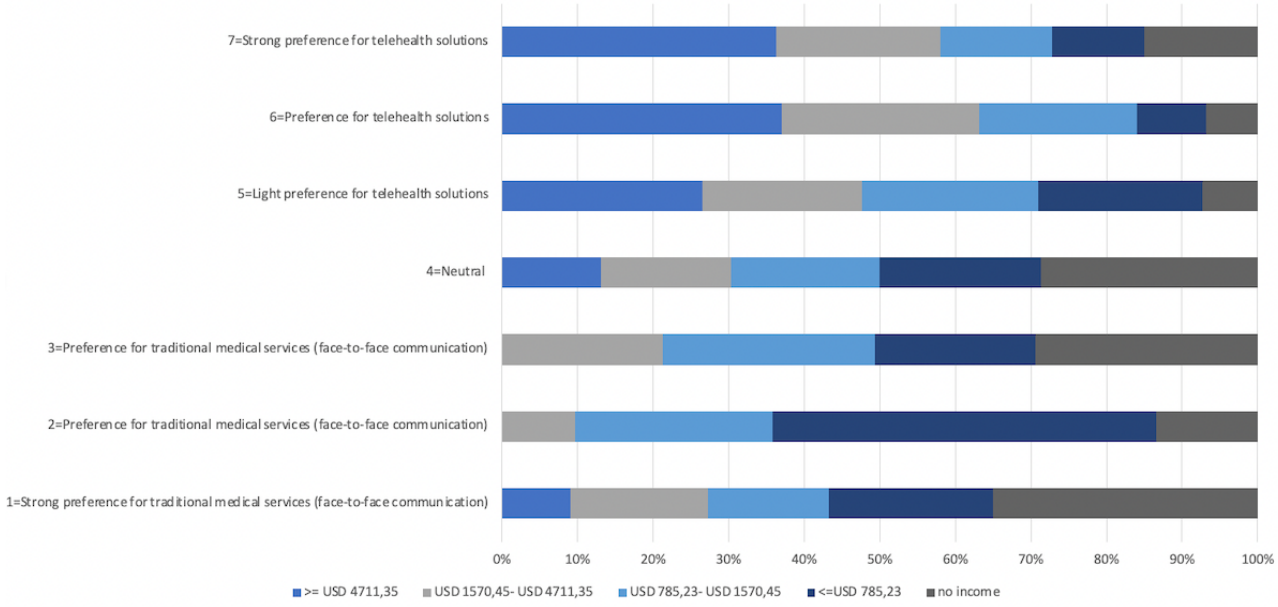
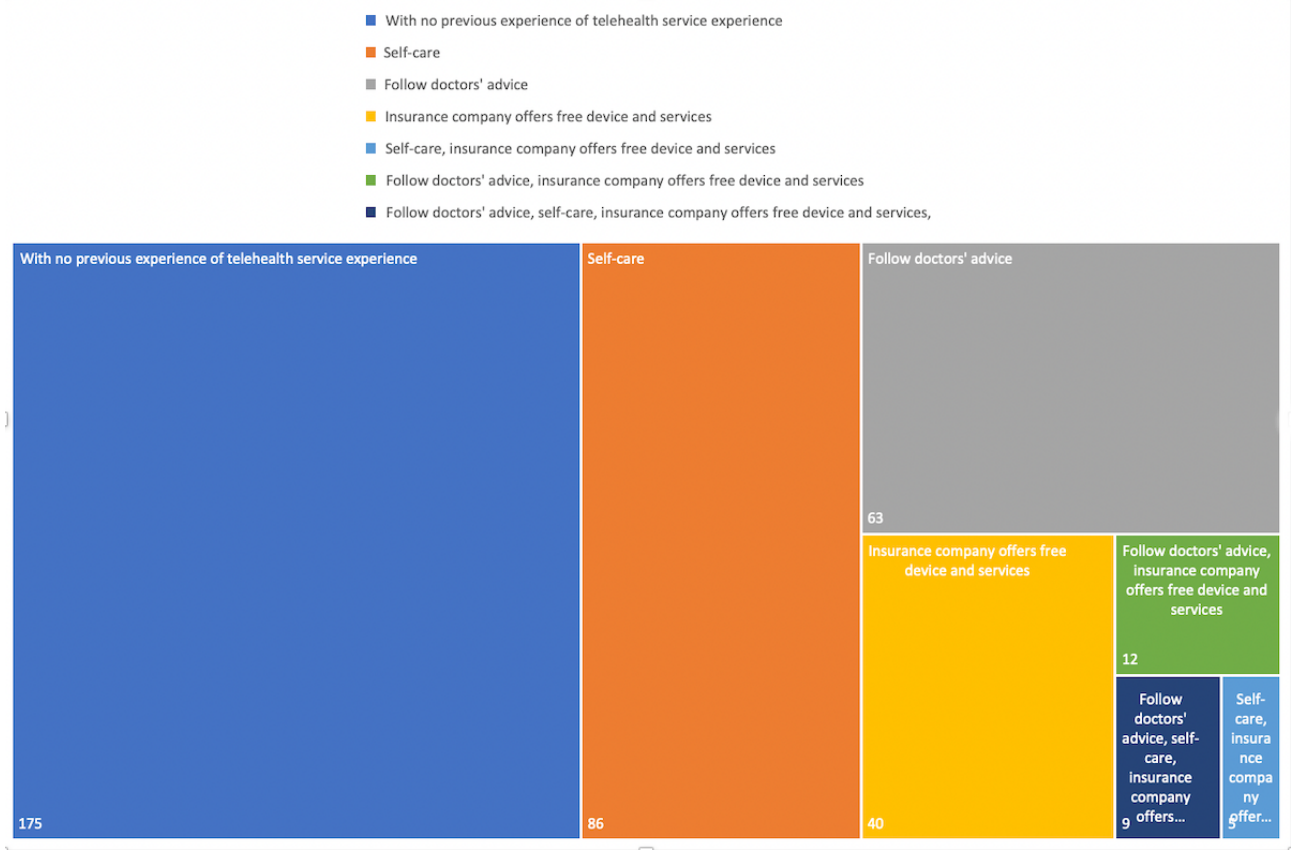


Figure 5. Reasons for using telehealth solutions.



The factors affecting the willingness to use telehealth solutions are ranked by the mean of each variable (Likert scale: 1-7, 1=no impact, 4=neutral, and 7=with an impact), as shown in Table 5. Among the 10 factors, 6 have means more than 4, suggesting that these factors influence the preference for telehealth solutions. The top 4 motivations are lowering health risks, raising health care awareness, lack of community medical services, and following the doctor's advice; these variables comprise Factor 1.

Factor 2 consists of the price, privacy risk, social influence, design and brand of the solution, and the participants' coverage with insurance plans. The mean value of these variables is close to neutral or less than 4, suggesting that survey participants in general do not believe that these factors influence their willingness to use telehealth solutions.

The accuracy of the data (Factor 3) collected through telehealth solutions is also a key factor. Compared with traditional medical instruments and equipment having the shortcoming of inaccurate

XSL•FO
RenderX

data reading, telehealth solutions collect more accurate health data. However, most doctors and hospitals still do not trust data collected from telehealth solutions and do not use these data sources as the basis for diagnosis or treatment. This makes it difficult for users to trust the devices used for collecting health data and monitoring health status.

Table 5. Ranking of factors affecting the willingness to use telehealth solutions among elderly users.

Factors	Ranking	Mean	SE
Lowering health risk	1	5.96	1.672
Raising health awareness	2	5.85	1.676
Lack of community health care service	3	5.77	1.721
Following doctors' prescriptions	4	5.27	2.032
Price of the solution	5	4.37	2.462
Data accuracy	6	4.07	2.314
Design of the solution	7	3.72	2.498
Privacy risk	8	3.70	2.342
Social influence	9	3.44	2.540
Free device offered by insurance companies	10	2.77	2.157

Modeling Process

This section presents the quantitative analytical results.

To avoid heterogeneity issues, the Kaiser–Meyer–Olkin (KMO) and Bartlett test was conducted to examine the correlation between the Likert scale variables; the KMO score of 0.796 suggests that the sample is adequate for factor analysis. Then, principal component analysis (PCA) was performed to reduce the dimensions of the model and the correlation between variables. With the factor loading for each factor confirmed, the Likert scale variables were then ranked based on the mean value of each variable. The next step was to test if demographic factors influenced the 3 factors identified by PCA. This was confirmed with analysis of variance (ANOVA) and *t* tests.

The modeling process started with a correlation matrix (Pearson correlation and Spearman rank correlation) to test if the data have multicollinearity (see Multimedia Appendix 2). Then, the O-logit model was run using Stata 16.0 along with the control variables. During the final modeling step, 10 participants were randomly selected to determine if the prediction preference scores matched the choices made by the participants.

The KMO and Bartlett test was conducted on 10 Likert scale factors related to survey participant preferences. The KMO coefficient is 0.796 (>0.5) with the Sig. value of 0.000 in the Bartlett sphere being less than 0.05, indicating that there is a certain degree of correlation among the 10 factors. Dimension reduction among the 10 factors was deemed necessary for further analysis.

Factor analysis is a commonly used dimensionality reduction method. PCA and the varimax right-angle rotation method were used to extract 3 principal factors. These 3 principal factors could explain 64.149% of the total variance, with the first, second, and third factors explaining 28.364%, 25.196%, and 10.589% of the total variance, respectively (see Tables 6 and 7). The variables selected had a factor loading greater than 0.5 (See Table 6).

Factor 1 consists of the price (0.812), design (0.738), impact of private insurance coverage (0.713), social influence (0.706), and privacy risk (0.612).

Factor 2, involving health-related motivations, consists of lowering health risk (0.864), raising health awareness (0.818), lack of community health care services (0.771), and following the doctor's advice or prescription (0.701).

Factor 3, related to the trust for telehealth solutions, consists of the data accuracy variable. Users' trust levels for telehealth solutions are influenced by whether data collected from wearables or medical devices at home are accepted by doctors and hospitals. Therefore, trust is influenced directly by the data accuracy of the solution (0.762).

ANOVA and *t* tests were conducted for assessing whether the relative importance of the 3 main factors from PCA analysis differ, depending on the city of residence, age, gender, education level, health status, income, living situation, regular social activity, and regular exercise. The results are shown in Table 8.

XSL•FO
RenderX

Table 6. Principal component analysis.

Component	Factor loading	Eigenvalue	Variance contribution rate	Cumulative contribution rate
Factor 1		2.836	28.364	28.364
Price	0.812			
Brand and design	0.738			
Private insurance coverage	0.713			
Social influence	0.706			
Privacy risk	0.612			
Factor 2: Health-related motivations		2.520	25.196	53.560
Lower health risk	0.864			
Raise health awareness	0.818			
Lack of community health care service	0.771			
Unstable doctor-patient relationship	0.701			
Factor 3: Trust		1.059	10.589	64.149
Data accuracy	0.762			

Table 7. Total variance explaineda.

Factor	Rotation sums of squared loadings		
	Total	% variance	Cumulative %
1	2.836	28.364	28.364
2	2.520	25.196	53.560
3	1.059	10.589	64.149

[a]Extraction method: PCA.

First, we assumed that the relative importance of factors varies with the residence city. The results of variance analysis support this hypothesis.

The first factor is related to the price, brand, and design associated with the telehealth solutions; factors such as social influence and private insurance coverage are also included.

Second, we assumed that age plays a significant role in determining user preference. In this study, survey participants were divided into 4 age groups, 51-60, 61-70, 71-80, and over 80 years. Considering that 71.79% (280/390) of all the survey participants have chronic diseases, users from this age group may consider the relevant health benefits such as raising health awareness, lowering health risk, and improving access to health care more than other age groups. ANOVA results suggest that the second factor varies with age (P=.088). The second factor mainly reflects the belief that telehealth solutions can raise health awareness, lower health risk, improve doctor-patient relationships and amend the gap related to the lack of community health care services.

Gender is also one of the key factors affecting user preference. The hypothesis is that male and female users have a perceived value, perceived risk, and perceived benefit associated with telehealth solutions. Considering the binary factor of gender, the t test could verify our hypothesis. The results show that the trust factor is significant at the level of 1%. This suggests that male and female survey participants differ in their trust on the data accuracy risk related to telehealth solutions.

The survey classifies users' education levels by years into five categories: primary school (1-6 years), high school (6-12 years), vocational school (12-15 years), college education (15-18 years), and postgraduation (>=18 years). Our hypothesis is that Factors 1, 2, and 3 differ across different education levels. However, the ANOVA results reject our hypothesis. With data suggesting that 84.4% of all survey participants have high school or primary school education, the conclusion is that survey participants with less than 15 years of education show no difference in Factors 1, 2, and 3.

The health status of survey participants is divided into four categories: self-reported healthy, suboptimal health status, with chronic disease (does not affect life quality), and with chronic disease (affects life quality). The variance analysis results support our hypothesis. The third factor, trust over data accuracy regarding telehealth solutions, is statistically significant and is affected by the health status of survey participants.

ANOVA results suggest that household income has a statistically significant effect on Factor 2, consisting of factors pertaining to health-related motivations. Families with high household incomes can bear the cost of using telehealth solutions, thereby benefiting from active self-health management. Users in the lower household income group pay more attention to factors such as the price of telehealth solutions, often ignoring the need

XSL•FO
RenderX

for active health management. Factor 2 varies among different income groups.

Trust over data accuracy regarding telehealth solutions (Factor 3) is also affected by whether survey participants live with their children or grandchildren. The survey participants' living situations are categorized as prefer living alone, prefer living with spouse, prefer living with children, and prefer living with grandchildren. Usually, it is the children and grandchildren living with their parents or grandparents who pay for telehealth solutions and teach their parents and grandparents to use such solutions. The elderly people thus benefit from living with their children or grandchildren and trust the telehealth solutions more than those who live alone or with spouses only.

Table 8. One-way analysis of variance and two-sample *t* test.

Hypothesis testing and variance analysis value	Factor 1	Factor 2	Factor 3
Residence city			
F value (*df1, df2*)	5.718 (3, 386)	2.245 (3, 386)	4.075 (3, 386)
P value	.001	.083	.007
Age			
F value *(df1, df2)*	0.467 (3, 386)	2.195 (3, 386)	0.172 (3, 386)
P value	.71	.088	.92
Gender			
F value *(df1, df2)*	0.074 (1, 388)	2.128 (1, 388)	7.570 (1, 388)
P value	.79	.15	.006
Education			
F value *(df1, df2)*	1.186 (4, 385)	0.180 (4, 385)	1.374 (4, 385)
P value	.32	.95	.24
Health condition			
F value *(df1, df2)*	1.494 (3, 386)	1.128 (3, 386)	3.468 (3, 386)
P value	.22	.34	.02
Income			
F value *(df1, df2)*	1.261 (4, 385)	4.109 (4, 385)	1.436 (4, 385)
P value	.29	.003	.22
Living situation			
F value *(df1, df2)*	1.216 (5, 384)	1.136 (5, 384)	2.665 (5, 384)
P value	.30	.34	.022
Regular social activity			
F value *(df1, df2)*	5.998 (1, 388)	2.508 (1, 388)	2.083 (1, 388)
P value	.015	.11	.15
Regular exercise			
F value *(df1, df2)*	4.726 (1, 388)	3.963 (1, 388)	0.605 (1, 388)
P value	.03	.047	.44

The *t* test results suggest that regular social activity has a statistically significant effect on Factor 1. Peer pressure from regular social interaction may encourage users to choose telehealth solutions over social influence, brand and design, and insurance plans. Survey participants with poor physical conditions often lack social activity and are subject to less social influence when it comes to using telehealth solutions.

Factors 1 and 2 also differ in terms of whether users exercise regularly. Survey participants exercising regularly are more health conscious and more willing to spend on telehealth solutions such as wearables and believe that telehealth solutions may raise health awareness, lower health risks, amend the gap in community health care and ensure health continuity by improving unstable doctor-patient relationships.

Factor 1

Ordered logit regression results suggest that Factor 1 has no statistically significant impact on the preference for telehealth solutions (See Table 9, rows 1 and 3, columns 1 and 4). Therefore, hypothesis 1 is rejected.

XSL•FO
RenderX

Table 9. Ordered logit regression.

	(1)	(2)	(3)	(4)
	y	y	y	y
Factor 1				
Coefficient	–0.0227 (*P*=.81)	__[a]	—	–0.0687 (*P*=.47)
t value (*df*)	–0.2449 (379)	—	—	–0.7299 (377)
Factor 2				
Coefficient	—	0.4628 (*P*<.001)	—	0.4821 (*P*<.001)
t value (*df*)	—	4.7735 (379)	—	4.9667 (377)
Factor 3				
Coefficient	—	—	–0.2554 (*P*=.007)	–0.2856 (*P*=.003)
t value (*df*)	—	—	–2.7108 (379)	–3.0028 (377)
Living city				
Coefficient	–0.0235 (*P*=.80)	–0.0463 (*P*=.61)	–0.0062 (*P*=.95)	–0.0201 (*P*=.83)
t value (*df*)	–0.2604 (379)	–0.5109 (379)	–0.0688 (379)	–0.2208 (377)
Age				
Coefficient	0.0151 (*P*=.88)	–0.0126 (*P*=.90)	–0.0018 (*P*<.001)	–0.0354 (*P*=.72)
t value (*df*)	0.1553 (379)	–0.1291 (379)	–0.0180 (379)	–0.3599 (377)
Gender				
Coefficient	0.2882 (*P*=.13)	0.3820 (*P*=.046)	0.2349 (*P*=.22)	0.3426 (*P*=.075)
t value (*df*)	1.5150 (379)	1.9995 (379)	1.2331 (379)	1.7801 (377)
Education				
Coefficient	0.1317 (*P*=.28)	0.1232 (*P*=.31)	0.1565 (*P*=.20)	0.1432 (*P*=.25)
t value (*df*)	1.0778 (379)	1.0082 (379)	1.2692 (379)	1.1548 (377)
Health status				
Coefficient	0.0743 (*P*=.48)	0.1365 (*P*=.20)	0.1105 (*P*=.30)	0.1881 (*P*=.084)
t value (*df*)	0.7034 (379)	1.2821 (379)	1.0352 (379)	1.7290 (377)
Income				
Coefficient	0.3707 (*P*=.001)	0.2919 (*P*=.01)	0.3908 (*P*=.001)	0.3124 (*P*=.006)
t value (*df*)	3.3153 (379)	2.5934 (379)	3.4734 (379)	2.7589 (377)
Living situation				
Coefficient	0.0545 (*P*=.66)	0.0392 (*P*=.75)	0.0446 (*P*=.72)	0.0176 (*P*=.89)
t value (*df*)	0.4400 (379)	0.3158 (379)	0.3595 (379)	0.1403 (377)
Regular socialization				
Coefficient	0.1871 (*P*=.069)	0.1714 (*P*=.097)	0.1745 (*P*=.09)	0.1477 (*P*=.15)
t value (*df*)	1.8191 (379)	1.6619 (379)	1.6979 (379)	1.4264 (377)
Regular exercise				
Coefficient	–0.0572 (*P*=.64)	–0.0865 (*P*=.48)	–0.0739 (*P*=.55)	–0.1112 (*P*=.37)
t value (*df*)	–0.4688 (379)	–0.7035 (379)	–0.6018 (379)	–0.8957 (377)
N	390	390	390	390
Pseudo R^2	0.0169	0.0322	0.0217	0.0384

[a]Not applicable.

XSL•FO
RenderX

Factor 2

Hypothesis 2 suggests the correlation between the willingness to use telehealth solutions and the health-related reasons. The regression coefficients of Factor 2 in Models 2 and 4 in Table 9 are positive at the significance level (*P*=.01). Hence, hypothesis 2 is valid.

Factor 3 and Preference

Hypothesis 3 assumes that data accuracy risk has a significant impact on the preference of elderly users. Table 9 suggests that the belief in the accuracy of the data collected by telehealth solutions is negatively related to the preference for telehealth solutions in Models 3 and 4. The cut variable suggests there are 7 categories of the dependant variable, where cut 1 puts the category at the lower end when y equals to 0.

To validate the model, 10 samples were randomly selected from the 390 participants to predict the probability of each participant's preference for telehealth solutions. The prediction made by the model is compared with the answer in the questionnaire for validation. Taking the first randomly selected sample as an example, the model suggests that the survey participant is most likely to choose 1 (1=preference for traditional health solutions, 4=neutral, and 7=preference for telehealth solutions). The model suggests that the users' preference for telehealth solutions is low; this is consistent with the users' actual choice (Y=1). Among the 10 selected samples, the preferences of 8 were successfully predicted, as observed in Table 10. With a prediction rate of 80%, the model is validated. The possibilities of the respondents choosing different categories of preferences (1-7) are denoted as p1 to p7, and the highest possibility (in italics) is the respondent's choice.

Table 10. Model validation[a].

No.	p1	p2	p3	p4	p5	p6	p7	Y	Prediction results
1	*0.336*	0.153	0.149	0.127	0.107	0.070	0.058	1	yes
2	*0.437*	0.158	0.136	0.104	0.079	0.049	0.039	1	yes
3	*0.174*	0.111	0.139	0.153	0.163	0.131	0.129	1	yes
4	0.127	0.088	0.121	0.148	*0.178*	0.161	0.177	7	no
5	0.135	0.093	0.124	0.150	*0.176*	0.155	0.167	7	no
6	0.052	0.042	0.067	0.101	0.164	0.212	*0.362*	7	yes
7	0.044	0.036	0.059	0.091	0.154	0.213	*0.402*	7	yes
8	0.130	0.090	0.122	0.148	*0.177*	0.159	0.173	5	yes
9	*0.202*	0.122	0.145	0.152	0.153	0.117	0.110	1	yes
10	0.076	0.058	0.088	0.124	0.179	0.199	*0.276*	7	yes

[a]p1 to p7: possibilities of the respondents choosing different categories of preferences.

Discussion

Principal Findings

Digitalization of the health care system has been rapidly accelerated by COVID-19. Because of social distancing and the highly communicable nature of the disease, the use of telehealth solutions grew exponentially, with expenditure on such solutions increasing as well. The high number of COVID-19 patients has consumed hospital and medical resources rapidly, depriving medical care for many patients with chronic diseases. The importance of using telemonitoring and telehealth solutions inside and outside the hospital setting has become more important than ever.

This study analyzed questionnaire data collected on factors related to the preference for telehealth solutions in Shenzhen, Hangzhou, Wuhan, and Yichang. The preference is related to the following factors: F1-perceived value of telehealth solutions related to product price, design, and social influence; F3-perceived risk for telehealth solutions related to trust over telehealth solutions; and F2-perceived benefits for telehealth solutions in self-care and health management. ANOVA and *t* tests were conducted to verify the influence of demographic variables on Factors 1, 2, and 3. The ordered logit regression results suggest that the perceived value (F1) has no significant impact on the preference for telehealth solutions, indicating the homogeneity of current consumer-facing telehealth solutions. Perceived benefits in self-care and health management (F2) are positively related with the preference for telehealth solutions, and F3 (trust over data accuracy) is negatively correlated with the preference for telehealth solutions.

The preference distribution graph (Figure 2) suggests that females are more conservative than males in their preference for telehealth solutions. This may be related to their income, health education, and tendency to socialize. This will make female users more willing to communicate with doctors face to face and use traditional health solutions.

Reasons behind the differences in the preference for telehealth solutions among residents from different cities can be explained based on infrastructure differences, such as the smart health initiative in Shenzhen and Wuhan and the concentration of hospitals and other medical resources in Wuhan. Shenzhen and Wuhan started early in their big data + health initiative, whereas other cities started later. The big data + health initiative provides the necessary digital infrastructure (hospital information system) for health system digitalization including the digitalization of hospitals and the connection of primary health service clinics.

XSL•FO
RenderX

The preference for telehealth solutions varies by income, and this may exist because communications with doctors are part of social activity; such social activities are strongly related to health education, social influence, and health insurance coverage. The lower income groups have less insurance coverage and less health education; they have a strong preference for seeing doctors in person and spending hours at hospitals because they have fewer other social activities. Higher income groups have better health education, wider health insurance coverage, and considerably less willingness to spend time at hospitals. Therefore, as they can afford telehealth solutions, based on the premise that telehealth solutions are not covered by the social (employee or resident) medical insurance schemes in China, the higher income groups are more willing to use telehealth solutions rather than wait at hospitals.

It is still necessary to address the data interoperability issue between hospitals and the primary health service clinics to ensure patients' care continuity; this may also help reduce the concentration of patients at hospitals and divert them back to primary care institutions. After all, doctors at level-3 hospitals have no time for helping outpatients to address their lifestyle problems once they leave the hospital. This gap leaves room for community care centers to step in and advise patients and monitor them regularly. The vacuum for community care centers can be filled in by telehealth solutions for patient self-care.

There are more telehealth solution providers in Shenzhen and Hangzhou, as well as high-quality hospitals. Doctors and nurses are more acceptable for telehealth solutions in large cities such as Shenzhen and Hangzhou. A wide variety of well-designed and affordable telehealth solutions are available in Shenzhen and Hangzhou. With the fast-paced lifestyle in these 2 cities, elderly (aged 50 years and above) users are more willing to use telehealth solutions. Considering the differences in disposable monthly income, it is more likely that residents in Shenzhen and Hangzhou enjoy the coverage of private insurance. Users with private health insurance coverage are more likely to believe that the coverage of telehealth solutions has an impact on the willingness to use telehealth solutions (whether positive or negative).

The homogeneity of DTC telehealth solutions can lead to indifference in users over factors such as the price, design, privacy risk, and brand and design of telehealth solutions. Currently, telehealth solution providers focus on providing heterogeneous solutions at lower prices. This lowers product profits and deters progress made in data interoperability and the acceptance of telehealth solutions by health care service providers. Some solution providers choose to cut core component configurations to reduce costs, thus failing to guarantee the quality of the solution. Some solution providers have actively marketed their products by offering installments and interest-free loans to attract users. Although promotion and marketing remain important, equipment manufacturers may consider improving the competitiveness of their products by promoting the medical value of their solutions, integrating the solution with the electronic health record system, providing noninvasive monitoring equipment, improving data accuracy, and providing privacy protection.

The findings of the study suggest that users do believe that telehealth solutions can improve health awareness, reduce health risk, mend the gap in community health care services, and improve care continuity, thus having a positive impact on the willingness to use telehealth solutions. Meanwhile, doctors' suggestions and prescriptions play a role in driving users to choose telehealth solutions over traditional health solutions as well. Elderly (aged 50 years and above) users have a strong demand for self-care and health management anytime and anywhere. Telehealth solutions that are easy to operate, carry, and understand can effectively meet the demand of elderly users for self-care and health management. Moreover, with the global need for qualified clinicians [20], the demand for telehealth solutions is expected to increase.

Compared with traditional health management methods, telehealth solutions offer convenient ways for maintaining health records and health care management at home. However, elderly users do not trust data collected through telehealth solutions. For example, elderly patients with hypertension prefer to go to the doctor to check their blood pressure instead of using the Bluetooth-connected blood pressure monitor at home. With the lack of integration of EHR systems, doctors cannot use the discontinuous data, even if the data collected by telehealth devices are relatively more accurate. There are also technical trust challenges regarding whether the algorithms are trained using accurate data representative of the potential user group. Human trust in the usability of the system, and the regulatory trust issues related to the ethical, legal, and social implications of the use of AI in health care are also important aspects to address [21]. With elderly users, it is extremely important to address the system usability issue of telehealth solutions and build human-level trust.

In the stakeholder interview stage, an interviewed doctor stated that elderly people with chronic diseases are more willing to go to the doctor for blood pressure measurement. The willingness to use telehealth solutions is affected by the users' health condition. For instance, telehealth solutions can provide users with a large amount of real-time personal health data, such as the heart rate, blood pressure, blood sugar, and other health indexes. For users with chronic diseases, although the data collected by telehealth solutions have certain reference values, doctors either have no access to the data or do not trust the data collected at home. With little or no integration with the health care system, elderly users do not trust the data collected through telehealth solutions.

Conclusions

Many existing telehealth solutions have similar features, prices, and designs. However, doctors often fail to use the data collected from telehealth solutions in making diagnosis and treatment decisions. They often use data collected from telehealth solutions as a reference but rarely base their decision on the daily data trends. Faced with increasing amounts of data from patients, resolving the data interoperability challenges between telehealth solution systems and hospital EHR systems seems more urgent than ever. Further research can focus on data interoperability between the EHR systems and telehealth solutions. The medical value of telehealth solutions can improve if doctors could

XSL•FO
RenderX

interpret data collected from telehealth solutions; furthermore, if doctors could diagnose, provide treatments, and adjust health care management plans based on such data, telehealth solutions can be included in insurance packages, making them more accessible.

There are hurdles to building trust for using telehealth solutions and AI in health care. Future research can also be extended to address such challenges by analyzing how to improve the transparency of algorithms by disclosing the data source and how the algorithms were built.

Owing to the limited scale of the questionnaire study (N=390), this paper only serves as a reference for exploring the implementation of telehealth solutions among elderly users in the next 5 to 10 years. The study was carried out in 4 Chinese cities (Shenzhen, Hangzhou, Wuhan, and Yichang), where the public health system differs in many ways from the health care systems in the United States or Europe. Therefore, the study is limited in scale with a sample size of 390 and remains an exploratory stage study. Further work can be done on the preference for telehealth solutions post-COVID and the changes in business models for telehealth solutions.

Acknowledgments

This work was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie–ITN Industrial Doctorate (grant 766139) and the National Key R&D Program of China (grant 2017YFE0112000). This article reflects only the author's views, and the European Research Executive Agency (REA) is not responsible for any use that may be made of the information it contains. The study received support from Dr. Gerd Spekowius from Philips Research, Dr. Lv Ping from University of China Academy of Sciences, Dr. Wei Chen from Fudan University, Dr. Francesca Spigarelli from University of Macerata, and Jerry Yu from Philips China.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English version of the questionnaire.
[PDF File (Adobe PDF File), 165 KB - mhealth_v10i1e27272_app1.pdf]

Multimedia Appendix 2
Pearson test and Spearman rank correlation test.
[DOCX File , 42 KB - mhealth_v10i1e27272_app2.docx]

References

1. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors determining patients' choice between mobile health and telemedicine: predictive analytics assessment. JMIR Mhealth Uhealth 2019 Jun;7(6):e13772 [FREE Full text] [doi: 10.2196/13772] [Medline: 31199332]
2. China's population to hit 'turning point' in 2026-2030 - think tank. Reuters. 2021 May 12. URL: https://www.reuters.com/world/china/chinas-population-hit-turning-point-2026-2030-think-tank-2021-05-12/ [accessed 2021-07-01]
3. At 54, China's average retirement age is too low. The Economist. 2021 Jun 25. URL: https://www.economist.com/china/2021/06/22/chinas-average-retirement-age-is-ridiculously-low-54 [accessed 2021-10-08]
4. Alexander GC, Tajanlangit M, Heyward J, Mansour O, Qato DM, Stafford RS. Use and content of primary care office-based vs telemedicine care visits during the COVID-19 pandemic in the US. JAMA Netw Open 2020 Oct;3(10):e2021476 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.21476] [Medline: 33006622]
5. Canhoto A, Arp S. Exploring the factors that support adoption and sustained use of health and fitness wearables. J Mark Manag 2016 Oct;33(1-2):32-60 [FREE Full text] [doi: 10.1080/0267257x.2016.1234505]
6. Fiebig D, Keane M, Louviere J, Wasi N. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Mark Sci 2010 May;29(3):393-421 [FREE Full text] [doi: 10.1287/mksc.1090.0508]
7. Welch BM, Harvey J, O'Connell NS, McElligott JT. Patient preferences for direct-to-consumer telemedicine services: a nationwide survey. BMC Health Serv Res 2017 Nov;17(1):784 [FREE Full text] [doi: 10.1186/s12913-017-2744-8] [Medline: 29183372]
8. Polinski JM, Barker T, Gagliano N, Sussman A, Brennan TA, Shrank WH. Patients' satisfaction with and preference for telehealth visits. J Gen Intern Med 2016 Mar;31(3):269-275 [FREE Full text] [doi: 10.1007/s11606-015-3489-x] [Medline: 26269131]
9. Minute Clinic. CVS Health. 2021. URL: https://cvshealth.com/our-services/health-and-wellness-services/minuteclinic [accessed 2021-07-15]
10. Miner H, Fatehi A, Ring D, Reichenberg JS. Clinician telemedicine perceptions during the COVID-19 pandemic. Telemed J E Health 2021 May;27(5):508-512. [doi: 10.1089/tmj.2020.0295] [Medline: 32946364]

XSL•FO
RenderX

11. Segrelles-Calvo G, Chiner E, Fernández-Fabrellas E. Acceptance of telemedicine among healthcare professionals. Arch Bronconeumol 2015 Dec;51(12):611-612. [doi: 10.1016/j.arbres.2015.08.002] [Medline: 26497416]
12. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
13. Burrus MT, Werner BC, Starman JS, Kurkis GM, Pierre JM, Diduch DR, et al. Patient perceptions and current trends in internet use by orthopedic outpatients. HSS J 2017 Oct;13(3):271-275 [FREE Full text] [doi: 10.1007/s11420-017-9568-2] [Medline: 28983221]
14. Marconi GP, Chang T, Pham PK, Grajower DN, Nager AL. Traditional nurse triage vs physician telepresence in a pediatric ED. Am J Emerg Med 2014 Apr;32(4):325-329 [FREE Full text] [doi: 10.1016/j.ajem.2013.12.032] [Medline: 24445223]
15. Baker LC, Johnson SJ, Macaulay D, Birnbaum H. Integrated telehealth and care management program for Medicare beneficiaries with chronic disease linked to savings. Health Aff (Millwood) 2011 Sep;30(9):1689-1697 [FREE Full text] [doi: 10.1377/hlthaff.2011.0216] [Medline: 21900660]
16. Kaambwa B, Ratcliffe J, Shulver W, Killington M, Taylor A, Crotty M, et al. Investigating the preferences of older people for telehealth as a new model of health care service delivery: a discrete choice experiment. J Telemed Telecare 2017 Feb;23(2):301-313. [doi: 10.1177/1357633X16637725] [Medline: 26985004]
17. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors determining patients' choice between mobile health and telemedicine: predictive analytics assessment. JMIR Mhealth Uhealth 2019 Jun;7(6):e13772 [FREE Full text] [doi: 10.2196/13772] [Medline: 31199332]
18. Sommer J, Dyczmons J, Grobosch S, Gontscharuk V, Vomhof M, Roden M, et al. Preferences of people with type 2 diabetes for telemedical lifestyle programmes in Germany: protocol of a discrete choice experiment. BMJ Open 2020 Sep;10(9):e036995 [FREE Full text] [doi: 10.1136/bmjopen-2020-036995] [Medline: 32907900]
19. Fletcher B, Hinton L, McManus R, Rivero-Arias O. Patient preferences for management of high blood pressure in the UK: a discrete choice experiment. Br J Gen Pract 2019 Sep;69(686):e629-e637 [FREE Full text] [doi: 10.3399/bjgp19X705101] [Medline: 31405832]
20. Kathy O'Reilly. Pat Baird, Head of Global Software Standards at Philips, discusses AI in healthcare at CES 2021. Philips. 2021 Jan 08. URL: https://tinyurl.com/4mw3mep2 [accessed 2021-07-15]
21. Ziegel ER, Neter J, Kutner M, Nachtsheim C, Wasserman W. Applied linear statistical models. Technometrics 1997 Aug;39(3):342. [doi: 10.2307/1271154]

Abbreviations

ANOVA: analysis of variance
DCE: discrete choice experiment
KMO: Kaiser–Meyer–Olkin
PCA: principal component analysis

Edited by C Dias; submitted 20.01.21; peer-reviewed by P Nissaisorakarn, B Chaudhry; comments to author 24.02.21; revised version received 16.07.21; accepted 27.10.21; published 12.01.22.

Please cite as:
Chen N, Liu P
Assessing Elderly User Preference for Telehealth Solutions in China: Exploratory Quantitative Study
JMIR Mhealth Uhealth 2022;10(1):e27272
URL: https://mhealth.jmir.org/2022/1/e27272
doi:10.2196/27272
PMID:35019852

©Nuoya Chen, Pengqi Liu. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 12.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Dose–Response Effects of Patient Engagement on Health Outcomes in an mHealth Intervention: Secondary Analysis of a Randomized Controlled Trial

Yiran Li[1*], MMSc; Yan Guo[1,2,3*], PhD; Y Alicia Hong[4], PhD; Yu Zeng[1], MMSc; Aliza Monroe-Wise[5], MD, MSc; Chengbo Zeng[6,7], PhD; Mengting Zhu[8], MS; Hanxi Zhang[9], PhD; Jiaying Qiao[1], MMSc; Zhimeng Xu[1], MMSc; Weiping Cai[10], MD; Linghua Li[10], MD; Cong Liu[10], MSN

[1]Department of Medical Statistics, School of Public Health, Guangzhou, China

[2]Sun Yat-sen Center for Global Health, Guangzhou, China

[3]Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

[4]Department of Health Administration and Policy, College of Health and Human Services, George Mason University, Fairfax, VA, United States

[5]Department of Global Health, University of Washington, Seattle, WA, United States

[6]South Carolina SmartState Center for Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[7]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[8]The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, Hong Kong

[9]National Center of AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[10]Department of Infectious Diseases, Guangzhou Eighth People's Hospital, Guangzhou Medical University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Yan Guo, PhD
Department of Medical Statistics
School of Public Health
74 Zhongshan 2nd Road
Guangzhou, 510080
China
Phone: 86 020 87333239
Email: Yan.Guo1@umassmed.edu

Abstract

Background: The dose–response relationship between patient engagement and long-term intervention effects in mobile health (mHealth) interventions are understudied. Studies exploring long-term and potentially changing relationships between patient engagement and health outcomes in mHealth interventions are needed.

Objective: This study aims to examine dose–response relationships between patient engagement and 3 psychosocial outcomes in an mHealth intervention, Run4Love, using repeated measurements of outcomes at baseline and 3, 6, and 9 months.

Methods: This study is a secondary analysis using longitudinal data from the Run4Love trial, a randomized controlled trial with 300 people living with HIV and elevated depressive symptoms to examine the effects of a 3-month mHealth intervention on reducing depressive symptoms and improving quality of life (QOL). We examined the relationships between patient engagement and depressive symptoms, QOL, and perceived stress in the intervention group (N=150) using 4–time-point outcome measurements. Patient engagement was assessed using the completion rate of course assignments and frequency of items completed. Cluster analysis was used to categorize patients into high- and low-engagement groups. Generalized linear mixed effects models were conducted to investigate the dose–response relationships between patient engagement and outcomes.

Results: The cluster analysis identified 2 clusters that were distinctively different from each other. The first cluster comprised 72 participants with good compliance to the intervention, completing an average of 74% (53/72) of intervention items (IQR 0.22). The second cluster comprised 78 participants with low compliance to the intervention, completing an average of 15% (11/72) of intervention items (IQR 0.23). Results of the generalized linear mixed effects models showed that, compared with the low-engagement group, the high-engagement group had a significant reduction in more depressive symptoms ($\beta=-1.93$; $P=.008$)

XSL•FO
RenderX

and perceived stress (β=−1.72; *P*<.001) and an improved QOL (β=2.41; *P*=.01) over 9 months. From baseline to 3, 6, and 9 months, the differences in depressive symptoms between the 2 engagement groups were 0.8, 1.6, 2.3, and 3.7 points, respectively, indicating widening between-group differences over time. Similarly, between-group differences in QOL and perceived stress increased over time (group differences in QOL: 0.9, 1.9, 4.7, and 5.1 points, respectively; group differences in the Perceived Stress Scale: 0.9, 1.4, 2.3, and 3.0 points, respectively).

Conclusions: This study revealed a positive long-term dose–response relationship between patient engagement and 3 psychosocial outcomes among people living with HIV and elevated depressive symptoms in an mHealth intervention over 9 months using 4 time-point repeat measurement data. The high- and low-engagement groups showed significant and widening differences in depressive symptoms, QOL, and perceived stress at the 3-, 6-, and 9-month follow-ups. Future mHealth interventions should improve patient engagement to achieve long-term and sustained intervention effects.

Trial Registration: Chinese Clinical Trial Registry ChiCTR-IPR-17012606; https://www.chictr.org.cn/showproj.aspx?proj=21019

(JMIR Mhealth Uhealth 2022;10(1):e25586) doi:10.2196/25586

KEYWORDS

mHealth; patient engagement; dose–response relationship; long-term effect; generalized linear mixed effects model

Introduction

Background

Mobile health (mHealth) interventions have gained increasing momentum in delivering easily accessible, patient-centered, individually tailored, and potentially cost-effective programs for a range of psychosocial disorders [1-3]. Previous studies have demonstrated the effectiveness of mHealth interventions in improving psychological outcomes [4,5]. For example, internet-based cognitive behavioral therapy (CBT) has been proven to be effective in treating depressive symptoms [4]. Another study found that an mHealth intervention with self-care strategies reduced depressive symptoms in people living with HIV [5]. However, few studies have explored the impact of patient engagement with mHealth interventions on long-term patient outcomes. Patient engagement is defined as the degree to which a patient adheres to an intervention [6]. Examining the impact of patient engagement on intervention effects beyond the termination of the intervention could help better understand the dose–response relationship in interventions.

Limited mHealth studies have examined the dose–response relationship in program evaluation. Of the few such studies, the most used is the pre–post design with short-term follow-up, typically within 3 months of the intervention [7-10]. To the best of our knowledge, there is only 1 mHealth study that aimed to explore the long-term dose–response relationship between patient engagement and mental health outcomes over 9 months [11]. This 3-month computerized CBT intervention found that different measures of patient engagement, such as number of log-ins, total time spent on the program, and number of visits to a mood diary (1 component of the intervention), were significantly associated with reduced depressive symptoms immediately after the intervention. In the long term, the completion rate of homework assignments was a significant predictor of reduced depressive symptoms at 9 months. However, this study only measured outcomes at the 3- and 9-month follow-ups, and the logistic regression used in the study could not reveal the likely changing dose–response relationship over time. Longitudinal studies with repeated measurements (≥3) may allow us to examine the changing relationship between patient engagement and intervention effects over time. A better understanding of the potential time-varying relationship between patient engagement and mHealth intervention effects is warranted in the long term.

Long-term dose–response relationships have been examined more thoroughly in face-to-face interventions than in mHealth interventions. In face-to-face CBT interventions, homework assignments are considered indispensable to the effect of psychotherapy. Homework is defined as structured, specific, and therapeutic activities that are routinely completed by the participants between sessions. Homework tasks might include self-monitoring of mood, thoughts and behaviors, behavioral activation, or specific cognitive and behavioral skills, such as breathing exercises [12,13]. Studies have found that homework compliance has a positive impact on psychosocial outcomes, such as reduction of anxiety or depressive symptoms at 6 or 12 months after treatment in face-to-face CBT interventions [12,13]. Compared with face-to-face interventions, mHealth interventions allow for easy and repeated access to intervention materials for participants long after the formal intervention period, which may result in long-term and sustained intervention effects. Understanding the potential impacts of long-term patient engagement with interventions and its associated outcomes is crucial for progress tracking, intervention refinement, and future scale-up for mHealth interventions.

Furthermore, patient engagement in mHealth interventions may be different from face-to-face interventions as the former captures more multi-faceted aspects of patient engagement, such as log-ins, completion rate, frequency of items completed, and time spent on the program [7,14]. Our previous study examined the associations between patient engagement and intervention outcomes at 3 months in the Run4Love program. We found that a higher completion rate and a greater frequency of completed items were associated with fewer depressive symptoms at 3 months [15]. These automated and multi-dimensional patient engagement data may provide important insights into intervention progress tracking and interpretation of intervention mechanisms. Given that mHealth tools allow for the capture of more multi-faceted factors of patient engagement, more studies to evaluate the impacts of patient engagement on health outcomes in mHealth interventions are needed [14,16,17].

XSL•FO
RenderX

Objectives

This study aims to examine the potential time-varying dose–response relationships between patient engagement and intervention effects in a randomized controlled trial of an mHealth intervention, the Run4Love program, and to fill gaps in the literature. The Run4Love trial aimed to reduce stress and depressive symptoms and to improve quality of life (QOL) among people who lived with HIV and were concurrently experiencing elevated depressive symptoms. We estimated the impact of patient engagement on depressive symptoms (the primary outcome of the intervention), QOL, and perceived stress at the 3-, 6-, and 9-month follow-ups. We hypothesize that better patient engagement in an mHealth intervention could lead to better and sustained health outcomes in the long term.

Methods

Overview

This study is a secondary analysis using data from the Run4Love trial, a parallel randomized controlled trial, to examine the effects of a WeChat (Tencent Holdings Limited)-based intervention on reducing depressive symptoms in people living with HIV and elevated depressive symptoms. The study design and primary results of the Run4Love trial have been published elsewhere [18,19]. The trial was registered in the Chinese Clinical Trial Registry (ChiCTR-IPR-17012606). The study protocol was approved by the Institutional Review Board of Sun Yat-sen University and has been published [19].

Participants and Procedure

A total of 300 people living with HIV and elevated depressive symptoms were recruited from the outpatient department of a large hospital designated for HIV treatment in Guangzhou, the third largest city in China, in 2017. The participants were recruited if they (1) were aged ≥18 years, (2) were HIV-seropositive, (3) had elevated depressive symptoms (measured using a Center for Epidemiologic Studies-Depression [CES-D] score of ≥16), (4) were active users of WeChat, and (5) were willing to provide hair samples (to measure cortisol as a biomarker of chronic stress). Participants were excluded if they were (1) currently on psychiatric or psychological treatment, (2) unable to finish the questionnaire, and (3) unable to engage in the intervention (read or listen to the materials on WeChat or engage in physical exercise because of medical or other reasons). The participants who met the eligibility criteria and were willing to participate completed a baseline survey and were randomized into the intervention or waitlist control group. A total of 150 participants in the intervention group received a 3-month Run4Love intervention and a 3-month booster session; the participants in the control group received a brochure on HIV-related nutrition in addition to usual care for HIV treatment. We used the data from the 150 participants in the intervention group in the analyses in this study.

Run4Love mHealth Intervention

The Run4Love intervention consisted of two components: adapted cognitive behavioral stress management (CBSM) courses and physical activity promotions [20]. We adapted the evidence-based CBSM courses on stress management and coping skills to the local context and modified them into 65 items in multimedia formats, including short articles, audio clips, and posters. These items were sent via a self-developed, enhanced WeChat platform for 3 months. In the booster session, 7 materials that were read or listened to the most during the intervention were selected and resent to the participants in the next 3 months after the intervention. The articles were approximately 1300 words and took approximately 5 minutes to read; the audio clips were 5-10 minutes in length; and the posters were pictures with motivational captions, which took <30 seconds to read. Physical activity promotions consisted of goal-setting and information on the guidance and benefits of regular exercise. The enhanced WeChat platform had added functions of automatic information sending, course completion tracking, and weekly personalized feedback. Participants in the intervention group received up to US $2 as financial incentives based on their course content completion via WeChat accounts on a weekly basis.

Measurement

Overview

This study collected data on individuals' sociodemographic characteristics, patient engagement, and psychosocial outcomes, including depressive symptoms, QOL, and perceived stress. Psychosocial outcomes were assessed at baseline and 3-, 6-, and 9-month follow-ups, collected by research staff using electronic questionnaires on a tablet. Data on patient engagement were collected automatically using the enhanced WeChat platform. Sociodemographic characteristics included age, gender, marital status, sexual orientation, and educational level.

Patient Engagement

Patient engagement was assessed through the patient's completion rate of course assignments and frequency of items completed as these 2 measurements were recommended as reliable measures of patient engagement in mHealth interventions targeting psychosocial outcomes [14] and proven in our previous study [15]. The results from our previous study revealed that these 2 measurements were significantly associated with reduced depressive symptoms at 3 months in the Run4Love intervention. Therefore, we grouped the participants according to these 2 measurements of patient engagement using cluster analysis. A total of 72 intervention items in the form of short articles, audio clips, and posters were delivered to the participants, of which 65 (90%) were sent during the 3-month intervention and 7 (10%) were resent 0 to 3 months after the intervention as a booster. The completion rate was calculated as the percentage of items completed out of 72 by a participant. Items that were clicked by the participants were regarded as completed. The frequency of items completed referred to the total number of times the items were read or listened to by a participant during the 3-month intervention and booster session 0-3 months after the intervention. For example, if a participant read 1 item sent during the 3-month intervention twice and the same item in the booster material twice, the frequency of items completed was counted as 4. As the participants were encouraged to practice the skills for stress management from the CBSM courses repeatedly, the frequency of items completed was used to capture the repetition aspect of patient engagement.

XSL•FO
RenderX

Both completion rate and frequency of items completed were automatically tracked by the enhanced mHealth platform. Good reliability of the composite measurement of patient engagement was shown in the study, and the Cronbach α was .97.

Depressive Symptoms

Depressive symptoms were measured using the CES-D scale, one of the most widely used self-reported questionnaires on depressive symptoms in China [21-23]. The CES-D scale consists of 20 items, such as *I felt depressed* and *I did not feel like eating; my appetite was poor*, and each item is rated on a 4-point Likert scale ranging from 0 (*rarely or none of the time*) to 3 (*most or all of the time*). The CES-D scores range from 0 to 60, with higher scores indicating a higher level of depressive symptoms and 16 being the cut-off point for possible clinical depressive symptoms [24]. Scores ranging from 16 to 20, 21 to 25, and 26 to 60 are considered mild, moderate, and severe depressive symptoms, respectively [25]. Good reliability of the CES-D score was shown in the study, and the Cronbach α at baseline and the 3-, 6-, and 9-month follow-ups was .77, .76, .84, and .83, respectively.

QOL Measurement

QOL was measured using the World Health Organization Quality of Life HIV short version (WHOQOL-HIV BREF), with 31 items assessing 6 domains: physical, psychological, level of independence, social relationships, environment, and beliefs [26]. Each domain comprises items rated on a 5-point Likert scale. The WHOQOL-HIV BREF scores range from 24 to 120, with higher scores indicating better QOL. The WHOQOL-HIV BREF has been widely used in the Chinese population with HIV and has shown good validity and reliability [27-29]. In this study, the Cronbach α for the WHOQOL-HIV BREF at baseline and the 3-, 6-, and 9-month follow-ups was .84, .91, .94, and .94, respectively.

Perceived Stress

Perceived stress was assessed using the 10-item Chinese version of the Perceived Stress Scale (PSS-10) [30]. The PSS-10 is the most widely used validated instrument for assessing the perception of stress in Chinese population [31-33]. It assesses the participants' feelings and thoughts in the previous month (eg, *How often have you been upset because of something that happened unexpectedly?*). The PSS-10 scores range from 0 to 40, with higher scores indicating higher levels of perceived stress. Scores ranging from 0 to 13, 14 to 26, and 27 to 40 are considered low, moderate, and high levels of perceived stress, respectively [34]. In this study, the Cronbach α for the PSS-10 at baseline and the 3-, 6-, and 9-month follow-ups was .67, .65, .69, and .65, respectively.

Statistical Analysis

First, descriptive statistics of demographic characteristics, baseline depressive symptoms, QOL, and perceived stress were presented. Continuous variables with normal distribution were described using mean and SD, and those with skewed distribution were described using median and IQR. Categorical variables were described using numbers and percentages.

Second, hierarchical clustering was used to classify participants into different groups based on the 2 metrics of patient engagement. Cluster analysis is an exploratory classification technique to group participants into different categories based on their similarities in specific metrics. This statistical method helps identify different engagement groups taking different dimensions into account. Hierarchical clustering iteratively merges smaller clusters into larger clusters. The hierarchical clustering procedure was as follows:

1. A similarity distance matrix was constructed by calculating the pairwise distance between different observations. Each observation was assigned to an individual cluster; therefore, each observation represented 1 cluster.
2. The 2 clusters *r* and *s* with a minimum distance from each other were identified.
3. Clusters *r* and *s* were merged, and *r* was replaced with the new cluster. Cluster *s* was deleted, and distances between the new cluster and each of the old clusters were computed.
4. Steps 2 and 3 were repeated until the total number of clusters was 2 [35].

Through clustering, we categorized participants into high- and low-engagement groups based on both measurements of completion rate and frequency of items completed, which could be used to evaluate the effects of different levels of patient engagement on health outcomes. To verify the results of the cluster analysis, we used Wilcoxon rank-sum tests to compare patient engagement between the high- and low-engagement groups along the 2 metrics. In addition, Wilcoxon rank-sum tests were used to examine whether patient outcomes were balanced between the 2 engagement groups at baseline.

Finally, generalized linear mixed effects models (GLMMs) with fixed effects of the engagement groups and time and including time as a random effect were conducted to estimate the trajectories of patient outcomes of depressive symptoms, QOL, and perceived stress, by the 2 engagement groups over time. The random effects of time represented interindividual varying time trends. This allowed for the estimation of variance in the outcomes within and among these time groups. The GLMM allows for the simultaneous analysis of repeated measures in a longitudinal design, thus providing a more accurate estimation of changes in outcomes over time. It also allows for the inclusion of cases with missing data [36,37], making it well-suited for longitudinal data that is likely to have missing values, such as in this study.

In total, 3 GLMMs were conducted to evaluate the relationships between levels of patient engagement and 3 health outcomes over time, adjusting for time and baseline characteristics. The dependent variables were depressive symptoms, QOL, and perceived stress measured repeatedly at baseline and 3, 6, and 9 months, whereas the independent variables were the patient engagement groups (high- and low-engagement group, the latter as reference) and the 4 time points (baseline and 3, 6, and 9 months, with baseline as reference). Baseline characteristics were included in the GLMMs as control variables, including age, gender, marital status, sexual orientation, and education. Only statistically significant characteristics were retained in the final models. The model estimate of the coefficient for

XSL•FO
RenderX

engagement groups represented how the trajectories of health outcomes differed between the 2 groups. For example, in the GLMM with CES-D as the dependent variable, a significant negative coefficient suggested that the high-engagement group had a reduction in more depressive symptoms over 9 months than the low-engagement group. Statistical significance was defined as $P<.05$. All analyses were conducted using SPSS (version 25; IBM).

Results

Baseline Characteristics

The Run4Love trial recruited 300 participants, and the data from the 150 participants in the intervention group were used in this study (Table 1). The mean age of the 150 participants was 28 years. Most were men (142/150, 94.7%), nonheterosexual (130/150, 86.7%), unmarried (132/150, 88%), and well-educated (98/150, 65.3%) with at least some college education. The mean scores for the CES-D, QOL, and perceived stress were 23.9 (SD 6.4), 77.4 (SD 9.0), and 20.0 (SD 4.4), respectively, at baseline. The average scores of the CES-D scale and perceived stress were at moderate levels. The average QOL score was comparable with that of other people living with HIV [27].

Table 1. Baseline characteristics and outcomes of the people living with HIV and elevated depressive symptoms in the intervention group (N=150).

Variables	Values
Age (years), mean (SD)	28.0 (5.8)
Gender, n (%)	
Male	142 (94.7)
Female	8 (5.3)
Sexual orientation, n (%)	
Heterosexual	20 (13.3)
Homosexual, bisexual, or uncertain	130 (86.7)
Education, n (%)	
High school or lower	52 (34.7)
More than high school	98 (65.3)
Marital status, n (%)	
Single, divorced, or widowed	132 (88)
Married	18 (12)
CES-D[a], mean (SD)	23.9 (6.4)
QOL[b], mean (SD)	77.4 (9.0)
PSS-10[c], mean (SD)	20.0 (4.4)

[a]CES-D: Center for Epidemiologic Studies-Depression.

[b]QOL: quality of life.

[c]PSS-10: Perceived Stress Scale.

Cluster Analysis

The cluster analysis identified 2 clusters that were distinctively different from each other (Table 2). The first cluster consisted of 72 participants with good compliance to the intervention program, completing an average of 74% (53/72) of intervention items (IQR 0.22) and 82 items (IQR 35.50) when accounting for repeated visits. Specifically, the high-engagement group completed an average of 77% (50/65) of intervention items (IQR 0.23) in the 3-month intervention and 43% (3/7) of items (IQR 0.57) in the booster session. The second cluster consisted of 78 participants with low compliance to the intervention program, completing an average of 15% (11/72) of intervention items (IQR 0.23) in the intervention program and 15 items (IQR 23.25) accounting for repeated visits. Specifically, the low-engagement group completed an average of 17% (11/65) of intervention items (IQR 0.25) in the 3-month intervention and 0 (IQR 1.00) in the booster session. The results of the Wilcoxon rank-sum tests confirmed the significant differences between the 2 cluster groups in both engagement measurements (completion rate and frequency of items completed; $P<.001$), with the high-engagement group having significantly better compliance in both engagement measurements than the low-engagement group, verifying the 2 distinct groups categorized through the cluster analysis. There were no significant group differences in depressive symptoms, QOL, and perceived stress at baseline (CES-D: $P=.54$; QOL: $P=.45$; and PSS-10: $P=.25$), indicating a balance in the outcomes between the 2 engagement groups at baseline.

XSL•FO
RenderX

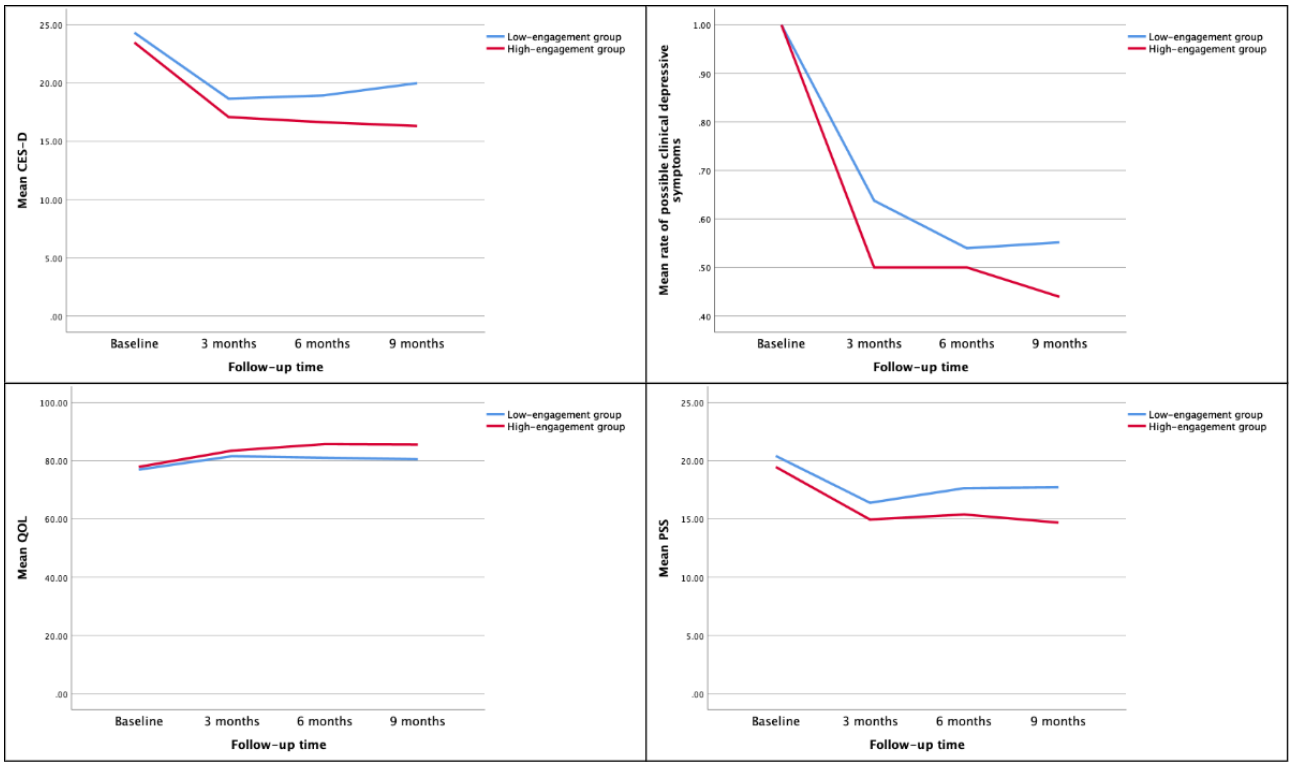
Table 2. Differences in the engagement measurements between the high- and low-engagement groups (N=150).

Engagement metrics	High-engagement group (n=72), median (IQR)	Low-engagement group (n=78), median (IQR)	Wilcoxon rank-sum test	*P* value
Completion rate				
3-month intervention	0.77 (0.23)	0.17 (0.25)	3091.0	<.001
Booster session	0.43 (0.57)	0.00 (0.14)	3877.5	<.001
Total	0.74 (0.22)	0.15 (0.23)	3091.5	<.001
Frequency of items completed				
3-month intervention	80 (30.50)	13 (21.25)	3084.0	<.001
Booster session	3 (4.75)	0 (1.00)	3955.5	<.001
Total	82 (35.50)	15 (23.25)	3084.5	<.001

Effects of Patient Engagement on Health Outcomes Over Time

The trajectories of the 3 outcomes at baseline and 3-, 6-, and 9-month follow-ups are shown in Figure 1, and GLMM results examining the effects of patient engagement on the 3 outcomes are presented in Table 3. Of the 150 participants in the intervention group, 139 (92.7%), 132 (88%), and 133 (88.7%) participants completed the follow-up surveys at 3, 6, and 9 months, respectively. These participants were randomly missing as there were no differences in demographic characteristics and outcomes at baseline between those who completed the follow-up surveys and those who did not. All 3 health outcomes in both groups significantly improved at 3 months immediately after the 3-month intervention (Figure 1). The results of the GLMM (Table 3) showed that the β coefficients of the 3 intervention outcomes were all statistically significant, indicating significant between-group differences in these outcomes over time.

Figure 1. Trajectories of depressive symptoms, quality of life (QOL), and perceived stress over time in high- and low-engagement groups. CES-D: Center for Epidemiologic Studies-Depression; PSS: Perceived Stress Scale.



XSL•FO
RenderX

Table 3. Effects of patient engagement on intervention outcomes at the 3-, 6-, and 9-month follow-ups: results from generalized linear mixed effects models[a].

Variables	β coefficient (SE; 95% CI)	*P* value
CES-D[b]		
Intercept	19.53 (1.86; 15.87 to 23.18)	<.001
Engagement group		
High vs low	−1.93 (0.72; −3.34 to −0.51)	.008
Follow-up		
3-month follow-up vs baseline	−6.02 (0.94; −7.87 to −4.16)	<.001
6-month follow-up vs baseline	−6.11 (0.99; −8.05 to −4.17)	<.001
9-month follow-up vs baseline	−5.78 (1.08; −7.91 to −3.65)	<.001
Age (years)	.16 (0.06; 0.03 to 0.28)	.01
Education		
High school or lower vs more than high school	2.83 (0.75; 1.35 to 4.31)	<.001
QOL[c]		
Intercept	78.55 (0.90; 76.77 to 80.33)	<.001
Engagement group		
High vs low	2.41 (0.93; 0.59 to 4.23)	.01
Follow-up		
3-month follow-up vs baseline	5.05 (1.22; 2.65 to 7.44)	<.001
6-month follow-up vs baseline	6.01 (1.27; 3.51 to 8.50)	<.001
9-month follow-up vs baseline	5.74 (1.34; 3.10 to 8.38)	<.001
Education		
High school or lower vs more than high school	−6.62 (0.97; −8.53 to −4.71)	<.001
PSS-10[d]		
Intercept	20.28 (0.45; 19.39 to 21.17)	<.001
Engagement group		
High vs low	−1.72 (0.45; −2.61 to −0.82)	<.001
Follow-up		
3-month follow-up vs baseline	−4.25 (0.60; −5.44 to −3.07)	<.001
6-month follow-up vs baseline	−3.42 (0.61; −4.63 to −2.22)	<.001
9-month follow-up vs baseline	−3.78 (0.65; −5.05 to −2.50)	<.001
Education		
High school or lower vs more than high school	1.51 (0.48; 0.57 to 2.44)	.002

[a]Generalized linear mixed effects models included all participants in the intervention group at all time points. Analyses were adjusted for individual characteristics, including age, gender, marital status, sexual orientation, and education at baseline, and those that did not show significant effects were removed from the final models. Educational level and age remained in the final models.

[b]CES-D: Center for Epidemiologic Studies-Depression.

[c]QOL: quality of life.

[d]PSS-10: Perceived Stress Scale.

The average group difference in CES-D scores between the high- and low-engagement groups was 1.93 (β=−1.93; *P*=.008) points over the 9 months, with the high-engagement group having lower levels of depressive symptoms than the low-engagement group. After the initial drastic decreases in both groups during the 3-month intervention (high-engagement group from 23.5 to 17.1 points; low-engagement group from 24.3 to 18.7 points), group differences in depressive symptoms increased in 0 to 6 months after the intervention. As shown in the first graph in Figure 1, there was a slowly rising trend in the CES-D scores at 3, 6, and 9 months (18.7, 18.9, and 20.0, respectively), suggesting some rebound in depressive symptoms

XSL•FO
RenderX

in the low-engagement group. In contrast, CES-D scores continued to decrease in the high-engagement group at 3, 6, and 9 months (17.1 to 16.6 and 16.3, respectively). Thus, between-group differences and standard effect sizes (Cohen *d*) in depressive symptoms (CES-D scores) increased over time, with 1.6, 2.3, and 3.7 points and 0.17, 0.23, and 0.34 points at 3, 6, and 9 months, respectively. The high-engagement group experienced a 14%, 4%, and 11% decrease in the rate of possible clinical depressive symptoms compared with the low-engagement group at 3, 6, and 9 months (50% vs 64%, 50% vs 54%, and 44% vs 55%, respectively).

Similar trends were also observed in the other 2 outcome measures. The average group differences in the QOL and perceived stress scores between the high- and low-engagement group were 2.41 (β=2.41; *P*=.01) and 1.72 (β=−1.72; *P*<.001) points over the 9 months, with the high-engagement group achieving better health outcomes across both measures. After the 3-month intervention, both outcome measures somewhat rebounded in the low-engagement group at 3, 6, and 9 months (QOL: 81.6, 81.0, and 80.5, respectively; PSS-10: 16.4, 17.7, and 17.7, respectively). In contrast, the high-engagement group had either continued improvement or reduced rebound effects at 3, 6, and 9 months in both outcome measures (QOL: 83.5, 85.7, and 85.6, respectively; PSS-10: 15.0, 15.4, and 14.7, respectively). Therefore, similar trends of widening between-group differences in QOL and perceived stress were observed over time at 3, 6, and 9 months (between-group differences in QOL: 1.9, 4.7, and 5.1 points, respectively; between-group differences in PSS-10: 1.4, 2.3, and 3.0 points, respectively). Standard effect sizes (Cohen *d*) in QOL (QOL scores) and perceived stress (PSS-10 scores) both increased over time, with 0.16, 0.37, and 0.39 in QOL and 0.25, 0.39, and 0.51 in PSS-10 at 3, 6, and 9 months, respectively.

Covariates including age, gender, marital status, sexual orientation, and educational level at baseline were adjusted in the 3 GLMMs, and only statistically significant variables were retained in the final models. Educational level was significantly associated with the 3 outcomes, and age was only significantly associated with depressive symptoms. Specifically, compared with those with higher education, participants with lower educational levels (high school or lower) had poorer health outcomes (CES-D: β=2.83, *P*<.001; QOL: β=−6.62, *P*<.001; PSS-10: β=1.51, *P*=.002). Compared with younger participants, older participants were more likely to report higher levels of depressive symptoms (β=.16; *P*=.01).

Discussion

Principal Findings

This study was among the first efforts to explore the potential time-varying dose–response relationships between patient engagement and various health outcomes over a span of 9 months using 4–time-point measurement data from the Run4Love mHealth intervention. The main finding was that patient engagement had a positive impact on the health outcomes, including depressive symptoms, QOL, and perceived stress, and such impacts were sustained over 9 months after the baseline. In addition, the dose–response relationship was not only sustained in the long term but also increased over time, as there were widening differences in health outcomes between the high- and low-engagement groups.

Both the high- and low-engagement groups benefited from the intervention, but the high-engagement group benefited more from the intervention consistently as the differences in health outcomes between the 2 groups became more pronounced over time. Such sustained and potentially increasing dose–response relationship in the long term has not been reported in previous mHealth studies. The increased effect sizes in depressive symptoms and QOL at 3, 6, and 9 months and increased effect sizes in perceived stress at 3 and 6 months were small to medium. The effect size of perceived stress at 9 months was medium [38]. In the low-engagement group, the intervention effects on health outcomes from 3 to 9 months are consistent with the findings of previous studies reporting rebound effects in mHealth or CBT interventions [39,40]. In contrast, the high-engagement group did not show similar trends; instead, it showed either sustained improvements or fewer rebound effects in all the health outcomes 0 to 6 months after the intervention, resulting in widening differences in these health outcomes between the 2 engagement groups.

Existing literature suggests that intervention effects tend to decrease or diminish over time after the intervention, with some demonstrating rebound effects [39,40]. For example, a face-to-face study found that women with breast cancer in the CBSM intervention group experienced a significant decrease in depressive symptoms immediately after treatment, but the level of depressive symptoms rebounded to baseline at the 1-month follow-up after the intervention [40].

The reasons for the sustained and widening dose–response relationship in this study are many; evidence-based interventions with rigorous design and good implementation are more likely to have sustained dose–response effects. What is missing in the literature on mHealth interventions is whether the momentum continues over time, for example, in 6- or 9-month follow-ups, and whether engagement level plays a role in this momentum. This study adds new evidence to this gap in the literature. Additional research to understand what factors predict patient engagement is also needed.

Our findings also revealed that education and age were important individual characteristics associated with the effects of the intervention, with participants with lower levels of education and of older age having poorer health outcomes. These findings are consistent with the literature [41-43]. To close the digital gap and bridge health disparities, mHealth interventions should be tailored to the needs of these more vulnerable groups, such as older people and those less educated. For example, mHealth interventions should be designed with easy-to-navigate interfaces, bigger fonts, and plain language with engaging multimedia such as pictures, audios, and videos [44].

Policy Implications

Given the sustained positive impacts of patient engagement on health outcomes found in this study, it is critical to improve patients' intervention adherence and engagement in both the intervention and subsequent booster sessions in mHealth

XSL•FO
RenderX

interventions. There are some effective ways to improve patient engagement as suggested in the literature and evidenced in our Run4Love trial. First, the intervention content needs to be culturally tailored and personalized, which requires formative research and a pilot study. The Run4Love intervention was developed based on extensive formative research and a pilot study [19,45]. We culturally adapted theory-guided and evidence-based CBSM courses, which have been proven effective in relieving depressive symptoms and improving other health outcomes in people who live with HIV [19,45]. Second, the program needs to have superior usability and user experience. Previous studies have shown that perceived usefulness and user-friendly experience are critical for improving mHealth engagement [46,47]. After many rounds of in-depth interviews with patients and the iterative development process, we designed and tailored the intervention platform and formats to meet their needs, such as delivering more appealing multimedia items [45]. With rigorous design and implementation, the participants reported high levels of satisfaction (92%-97%) at all 3 follow-ups in the Run4Love trial [18].

During the intervention, another way to improve patient engagement is to provide timely and personalized feedback. The enhanced WeChat platform automatically sends weekly feedback on the completion status to each participant. In addition to automatic weekly feedback, the Run4Love program also consisted of 5 phone calls made by the research staff at 1 week and 1, 2, 5, and 8 months to address technical challenges and motivate their participation. In addition, the backend platforms of mHealth interventions and wearable devices allow for the collection of passive data on various dimensions of patient engagement, such as physical activities, sleep hours and quality, log-in times, and time and duration of reading, listening to, or watching the intervention items [48-50]. Taking advantage of the easy-to-track user engagement data available in mHealth interventions is critical for process monitoring and quality control of the trials.

Limitations

There are several limitations to this study. First, the participants in this study were mostly from urban areas, young, and well-educated, and most were nonheterosexual men. Therefore, generalization of the results should be applied with caution. Second, measurement biases may exist in patient engagement. This study only measured patient engagement in CBSM courses but not in physical activity promotions as patient engagement data were not recorded or available in physical activities. However, from the qualitative interviews, we found that most of the engaged participants had a higher engagement in both CBSM courses and physical activities. Therefore, a long-term dose–response relationship between patient engagement in physical activities and intervention outcomes may exist, which needs to be further explored in future research. Moreover, the intervention items were considered as completed when clicked; therefore, we were not able to verify the actual completion or quality of completion. Nevertheless, the patient engagement metrics in this study served as a reliable measurement to assess the dose–response relationship, and the effect of such potential overestimation of patient engagement might only have diluted the observed dose–response relationship [15,51]. Finally, patient engagement in different formats of the intervention content, including short articles, audio clips, and posters, was not differentiated as this was beyond the objectives of this study. Future studies could further explore better measurements of patient engagement in terms of different intervention components and types of multimedia materials in mHealth interventions.

Conclusions

In conclusion, this study revealed a positive long-term dose–response relationship between patient engagement and 3 psychosocial outcomes in an mHealth intervention using 4–time-point measurement data over 9 months. High- and low-engagement groups showed significant and widening differences in depressive symptoms, QOL, and perceived stress at the 3-, 6-, and 9-month follow-ups in the Run4Love trial. Future mHealth interventions should improve patient engagement to achieve long-term and sustained intervention effects.

Acknowledgments

This study was supported by a grant from the National Natural Science Foundation of China (grant 71573290) and the China Medical Board (grant 17-271). The funder had no role in the study design, data collection and analysis, or writing of the manuscript.

Authors' Contributions

YL analyzed the data and drafted the manuscript. YG and YAH contributed to obtaining funding, the study design, and manuscript revision. YZ helped with the study concept. AMW helped with manuscript revision. CZ, MZ, HZ, JQ, and ZX contributed to the clinical trial and data acquisition. WC, LL, and CL provided administrative, technical, and material support for the clinical trial.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).

XSL•FO
RenderX

[PDF File (Adobe PDF File), 1687 KB - mhealth_v10i1e25586_app1.pdf]

References

1. Merchant R, Torous J, Rodriguez-Villa E, Naslund JA. Digital technology for management of severe mental disorders in low-income and middle-income countries. Curr Opin Psychiatry 2020 Sep;33(5):501-507. [doi: 10.1097/YCO.0000000000000626] [Medline: 32520747]
2. Bush NE, Armstrong CM, Hoyt TV. Smartphone apps for psychological health: a brief state of the science review. Psychol Serv 2019 May;16(2):188-195. [doi: 10.1037/ser0000286] [Medline: 30407057]
3. Marzano L, Bardill A, Fields B, Herd K, Veale D, Grey N, et al. The application of mHealth to mental health: opportunities and challenges. Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/S2215-0366(15)00268-0] [Medline: 26462228]
4. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 01;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]
5. Schnall R, Cho H, Mangone A, Pichon A, Jia H. Mobile health technology for improving symptom management in low income persons living with HIV. AIDS Behav 2018 Jan 03. [doi: 10.1007/s10461-017-2014-0] [Medline: 29299790]
6. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
7. Sanatkar S, Baldwin PA, Huckvale K, Clarke J, Christensen H, Harvey S, et al. Using cluster analysis to explore engagement and e-Attainment as emergent behavior in electronic mental health. J Med Internet Res 2019 Nov 28;21(11):e14728 [FREE Full text] [doi: 10.2196/14728] [Medline: 31778115]
8. Enrique A, Palacios JE, Ryan H, Richards D. Exploring the relationship between usage and outcomes of an internet-based intervention for individuals with depressive symptoms: secondary analysis of data from a randomized controlled trial. J Med Internet Res 2019 Aug 01;21(8):e12775 [FREE Full text] [doi: 10.2196/12775] [Medline: 31373272]
9. Geramita EM, Belnap BH, Abebe KZ, Rothenberger SD, Rotondi AJ, Rollman BL. The association between increased levels of patient engagement with an internet support group and improved mental health outcomes at 6-month follow-up: post-hoc analyses from a randomized controlled trial. J Med Internet Res 2018 Jul 17;20(7):e10402 [FREE Full text] [doi: 10.2196/10402] [Medline: 30021711]
10. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, et al. The association between adherence and outcome in an Internet intervention for depression. J Affect Disord 2018 Dec 15;229:443-449. [doi: 10.1016/j.jad.2017.12.028] [Medline: 29331706]
11. de Graaf LE, Huibers MJ, Riper H, Gerhards SA, Arntz A. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. J Affect Disord 2009 Aug;116(3):227-231. [doi: 10.1016/j.jad.2008.12.009] [Medline: 19167094]
12. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. Behav Ther 2016 Sep;47(5):755-772. [doi: 10.1016/j.beth.2016.05.002] [Medline: 27816086]
13. Glenn D, Golinelli D, Rose RD, Roy-Byrne P, Stein MB, Sullivan G, et al. Who gets the most out of cognitive behavioral therapy for anxiety disorders? The role of treatment dose and patient engagement. J Consult Clin Psychol 2013 Aug;81(4):639-649 [FREE Full text] [doi: 10.1037/a0033403] [Medline: 23750465]
14. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
15. Zeng Y, Guo Y, Li L, Hong YA, Li Y, Zhu M, et al. Relationship between patient engagement and depressive symptoms among people living with hiv in a mobile health intervention: secondary analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e20847 [FREE Full text] [doi: 10.2196/20847] [Medline: 33118956]
16. Baltierra NB, Muessig KE, Pike EC, LeGrand S, Bull SS, Hightow-Weidman LB. More than just tracking time: complex measures of user engagement with an internet-based health promotion intervention. J Biomed Inform 2016 Mar;59:299-307 [FREE Full text] [doi: 10.1016/j.jbi.2015.12.015] [Medline: 26732997]
17. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2016 Dec 13. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
18. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. J Med Internet Res 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: 10.2196/16715] [Medline: 32044751]
19. Guo Y, Hong YA, Qiao J, Xu Z, Zhang H, Zeng C, et al. Run4Love, a mHealth (WeChat-based) intervention to improve mental health of people living with HIV: a randomized controlled trial protocol. BMC Public Health 2018 Dec 26;18(1):793 [FREE Full text] [doi: 10.1186/s12889-018-5693-1] [Medline: 29940921]
20. Antoni M, Ironson G, Schneiderman N. Cognitive-Behavioral Stress Management for Individuals Living with HIV: Facilitator Guide. Oxford, United Kingdom: Oxford University Press; 2007:1-155.

XSL•FO
RenderX

21. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Measur 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
22. Wang M, Armour C, Wu Y, Ren F, Zhu X, Yao S. Factor structure of the CES-D and measurement invariance across gender in Mainland Chinese adolescents. J Clin Psychol 2013 Sep;69(9):966-979. [doi: 10.1002/jclp.21978] [Medline: 23775279]
23. Zhang Y, Ting RZ, Lam MH, Lam S, Yeung RO, Nan H, et al. Measuring depression with CES-D in Chinese patients with type 2 diabetes: the validity and its comparison to PHQ-9. BMC Psychiatry 2015 Aug 18;15:198 [FREE Full text] [doi: 10.1186/s12888-015-0580-0] [Medline: 26281832]
24. Zich JM, Attkisson CC, Greenfield TK. Screening for depression in primary care clinics: the CES-D and the BDI. Int J Psychiatry Med 1990;20(3):259-277. [doi: 10.2190/LYKR-7VHP-YJEM-MKM2] [Medline: 2265888]
25. Chwastiak L, Ehde DM, Gibbons LE, Sullivan M, Bowen JD, Kraft GH. Depressive symptoms and severity of illness in multiple sclerosis: epidemiologic study of a large community sample. Am J Psychiatry 2002 Nov;159(11):1862-1868. [doi: 10.1176/appi.ajp.159.11.1862] [Medline: 12411220]
26. Hsiung P, Fang C, Wu C, Sheng W, Chen S, Wang J, et al. Validation of the WHOQOL-HIV BREF among HIV-infected patients in Taiwan. AIDS Care 2011 Aug;23(8):1035-1042. [doi: 10.1080/09540121.2010.543881] [Medline: 21500023]
27. Zhu Y, Liu J, Qu B. Psychometric properties of the Chinese version of the WHOQOL-HIV BREF to assess quality of life among people living with HIV/AIDS: a cross-sectional study. BMJ Open 2017 Aug 21;7(8):e016382 [FREE Full text] [doi: 10.1136/bmjopen-2017-016382] [Medline: 28827253]
28. Shan D, Ge Z, Ming S, Wang L, Sante M, He W, et al. Quality of life and related factors among HIV-positive spouses from serodiscordant couples under antiretroviral therapy in Henan Province, China. PLoS One 2011;6(6):e21839 [FREE Full text] [doi: 10.1371/journal.pone.0021839] [Medline: 21738797]
29. Ming Z, Prybylski D, Cheng F, Airawanwat R, Zhu Q, Liu W, et al. Two-year prospective cohort study on quality of life outcomes among people living with HIV after initiation of antiretroviral therapy in Guangxi, China. J Assoc Nurses AIDS Care 2014;25(6):603-613. [doi: 10.1016/j.jana.2014.04.003] [Medline: 24950656]
30. Lu W, Bian Q, Wang W, Wu X, Wang Z, Zhao M. Chinese version of the Perceived Stress Scale-10: a psychometric study in Chinese university students. PLoS One 2017;12(12):e0189543 [FREE Full text] [doi: 10.1371/journal.pone.0189543] [Medline: 29252989]
31. Leung DY, Lam T, Chan SS. Three versions of Perceived Stress Scale: validation in a sample of Chinese cardiac patients who smoke. BMC Public Health 2010 Aug 25;10:513 [FREE Full text] [doi: 10.1186/1471-2458-10-513] [Medline: 20735860]
32. Ng S. Validation of the 10-item Chinese perceived stress scale in elderly service workers: one-factor versus two-factor structure. BMC Psychol 2013;1(1):9 [FREE Full text] [doi: 10.1186/2050-7283-1-9] [Medline: 25566361]
33. Wang Z, Chen J, Boyd JE, Zhang H, Jia X, Qiu J, et al. Psychometric properties of the Chinese version of the Perceived Stress Scale in policewomen. PLoS One 2011;6(12):e28610. [doi: 10.1371/journal.pone.0028610] [Medline: 22164311]
34. Olpin M, Hesson M. Stress Management for Life: A Research-Based Experiential Approach. Toronto, Canada: Nelson Education; 2015:1-432.
35. Frades I, Matthiesen R. Overview on techniques in cluster analysis. Methods Mol Biol 2010;593:81-107. [doi: 10.1007/978-1-60327-194-3_5] [Medline: 19957146]
36. McCulloch CE, Neuhaus JM. Generalized linear mixed models. In: Wiley StatsRef: Statistics Reference Online. Hoboken, New Jersey, United States: John Wiley & Sons, Ltd; 2014.
37. Hadfield JD. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R Package. J Stat Soft 2010;33(2):1-22. [doi: 10.18637/jss.v033.i02]
38. Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol 2005 Dec;58(12):1217-1219. [doi: 10.1016/j.jclinepi.2005.07.009] [Medline: 16291464]
39. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
40. Beatty L, Koczwara B. An effectiveness study of a CBT group program for women with breast cancer. Clin Psychol 2010 Jul 20;14(2):45-53. [doi: 10.1080/13284207.2010.500307]
41. Hong YA, Zhou Z, Fang Y, Shi L. The digital divide and health disparities in China: evidence from a national survey and policy implications. J Med Internet Res 2017 Sep 11;19(9):e317 [FREE Full text] [doi: 10.2196/jmir.7786] [Medline: 28893724]
42. DiMaggio P, Hargittai E, Celeste C, Shafer S. From unequal access to differentiated use: A literature review and agenda for research on digital inequality. Russel Sage Foundation. 2004. URL: http://www.eszter.com/research/pubs/dimaggio-etal-digitalinequality.pdf [accessed 2021-12-14]
43. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
44. Parker SJ, Jessel S, Richardson JE, Reid MC. Older adults are mobile too!Identifying the barriers and facilitators to older adults' use of mHealth for pain management. BMC Geriatr 2013;13:43 [FREE Full text] [doi: 10.1186/1471-2318-13-43] [Medline: 23647949]

XSL•FO
RenderX

45. Guo Y, Xu Z, Qiao J, Hong YA, Zhang H, Zeng C, et al. Development and feasibility testing of an mHealth (Text Message and WeChat) intervention to improve the medication adherence and quality of life of people living with HIV in China: pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 Sep 04;6(9):e10274 [FREE Full text] [doi: 10.2196/10274] [Medline: 30181109]
46. Hong Y, Goldberg D, Dahlke DV, Ory MG, Cargill JS, Coughlin R, et al. Testing usability and acceptability of a web application to promote physical activity (iCanFit) among older adults. JMIR Hum Factors 2014 Oct 13;1(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.3787] [Medline: 27025254]
47. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2015 Jan;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
48. Roy S, Srivastava S, Gururajan R. Integrating wearable devices and recommendation system: towards a next generation healthcare service delivery. J Inf Technol Theory Appl 2019;19(4):2 [FREE Full text]
49. Dinh-Le C, Chuang R, Chokshi S, Mann D. Wearable health technology and electronic health record integration: scoping review and future. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e12861 [FREE Full text] [doi: 10.2196/12861] [Medline: 31512582]
50. Torous J, Kiang MV, Lorme J, Onnela J. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research directions. JMIR Ment Health 2016;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
51. Beintner I, Vollert B, Zarski A, Bolinski F, Musiat P, Görlich D, et al. Adherence reporting in randomized controlled trials examining manualized multisession online interventions: systematic review of practices and proposal for reporting standards. J Med Internet Res 2019 Aug 15;21(8):e14181 [FREE Full text] [doi: 10.2196/14181] [Medline: 31414664]

Abbreviations

CBSM: cognitive behavioral stress management
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies-Depression
GLMM: generalized linear mixed effects model
mHealth: mobile health
PSS-10: Perceived Stress Scale
QOL: quality of life
WHOQOL-HIV BREF: World Health Organization Quality of Life HIV short version

Edited by L Buis; submitted 08.11.20; peer-reviewed by E White, J Villegas, S Wang, K Fuji; comments to author 15.03.21; revised version received 04.05.21; accepted 12.11.21; published 04.01.22.

Please cite as:
Li Y, Guo Y, Hong YA, Zeng Y, Monroe-Wise A, Zeng C, Zhu M, Zhang H, Qiao J, Xu Z, Cai W, Li L, Liu C
Dose–Response Effects of Patient Engagement on Health Outcomes in an mHealth Intervention: Secondary Analysis of a Randomized Controlled Trial
JMIR Mhealth Uhealth 2022;10(1):e25586
URL: https://mhealth.jmir.org/2022/1/e25586
doi:10.2196/25586
PMID:34982724

©Yiran Li, Yan Guo, Y Alicia Hong, Yu Zeng, Aliza Monroe-Wise, Chengbo Zeng, Mengting Zhu, Hanxi Zhang, Jiaying Qiao, Zhimeng Xu, Weiping Cai, Linghua Li, Cong Liu. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 04.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effect of an Integrative Mobile Health Intervention in Patients With Hypertension and Diabetes: Crossover Study

Sang Woo Oh[1], MD, PhD; Kyoung-Kon Kim[2], MD, PhD; Sung Soo Kim[3], MD, PhD; Su Kyung Park[1], MSDH; Sangshin Park[4], DVM, PhD

[1]Department of Family Medicine, Dongguk University Ilsan Hospital, Dongguk University College of Medicine, Gyeonggi-do, Republic of Korea

[2]Department of Family Medicine, Gachon University Gil Medical Center, Gachon University College of Medicine, Incheon, Republic of Korea

[3]Department of Family Medicine, Chungnam National University Hospital, Chungnam National University College of Medicine, Daejeon, Republic of Korea

[4]Graduate School of Urban Public Health & Department of Urban Big Data Convergence, University of Seoul, Seoul, Republic of Korea

Corresponding Author:
Sang Woo Oh, MD, PhD
Department of Family Medicine
Dongguk University Ilsan Hospital
Dongguk University College of Medicine
27, Dongguk-ro, Ilsandong-gu
Goyang-si
Gyeonggi-do, 10326
Republic of Korea
Phone: 82 31 961 7490
Email: osw6021@naver.com

Abstract

Background: Obesity, hypertension, and type 2 diabetes mellitus (T2DM) are worldwide epidemics that inflict burdens on both public health and health care costs. Self-management plays an important role in the proper management of these 3 chronic diseases, and in this context, mobile health (mHealth) can be a cost-effective self-management tool.

Objective: The aim of this pilot study is to evaluate the effects of an integrative mHealth approach for obesity, hypertension, and T2DM on body fat, blood pressure, and blood glucose levels and demonstrate the clinical outcomes. The participants were patients aged 40 to 70 years who were treated for T2DM (hemoglobin A_{1c} [HbA_{1c}] above 6.0%) without insulin or hypertension and obesity, controlled with pharmacotherapy.

Methods: This pilot study was performed using a controlled, randomized, 3-month, 2-period crossover design. A total of 37 participants were recruited from 2 university hospitals in South Korea. Integrative mHealth comprised 4 parts: self-measuring home devices for monitoring blood glucose and blood pressure; 2 smartphone apps, where one gathered lifestyle data, giving them feedback with health information, and the other provided drug information and reminders of the medication schedule; unmanned kiosks for official measurement of blood pressure and body composition; and web-based access to participants' health information.

Results: Data from the 32 participants were analyzed. Their mean HbA_{1c} level was 7.5% (SD 0.8, ranging from 6.1% to 9.4%). Approximately 38% (12/32) of the participants had hypertension. BMIs of all participants except 1 were >23 kg/m^2. The input rates of food intake and exercise to the smartphone app were very low (24.9% and 5.3%, respectively). On the contrary, the input rate of medicine intake was high (84.0%). Moreover, there was no significant difference in the input rate of taking medicine irrespective of whether the mHealth period was before or after the conventional treatment period (80.3% and 87.3%, respectively; P=.06). Among the 3 input functions of food intake, exercise, and medicine intake in smartphone apps, the input of medicine intake was a more helpful, easier to use, and better-designed function than the others. There were no significant differences in changes in body weight (−0.519 kg vs 0 kg), BMI (−0.133 kg/m^2 vs −0.167 kg/m^2), body composition (body fat −0.255% vs 0.172%), blood pressure (systolic −0.226 mm Hg vs −2.839 mm Hg), and HbA_{1c} (−0.269% vs –0.009%) between the integrative mHealth and conventional treatment groups. However, in proportion to the elevation in the input rate of taking medicine, body fat mass (P=.04) and HbA_{1c} (P=.03) were lower in the integrative mHealth group.

XSL•FO
RenderX

Conclusions: Although smartphone apps can influence body fat and blood glucose levels, they have failed to show clinical improvement. A higher input rate of taking medicine was related to significantly lower body fat mass and HbA_{1c} levels.

(JMIR Mhealth Uhealth 2022;10(1):e27192) doi:10.2196/27192

KEYWORDS

diabetes mellitus type 2; obesity; hypertension; mHealth; mobile phone

Introduction

Background

Obesity is the established main cause of hypertension and type 2 diabetes mellitus (T2DM) and can lead to the development of coronary vascular disease; furthermore, these 3 illnesses have attained the status of global epidemics recently [1-3]. Consequently, they inflict a huge economic burden on public health systems worldwide. Obesity was estimated to account for 0.7% to 2.8% of the total health care expenditure (HCE), and people with obesity had health care costs that were approximately 30% greater than those with normal weight [4]. The incremental medical expenditure ratios for people in Korea with BMIs of 30 kg/m^2 to 34.99 kg/m^2 and >35 kg/m^2 were 34.3% and 38.4%, respectively, as compared that of with people with BMIs of 18.5 kg/m^2 to 22.99 kg/m^2 from 2002 to 2013 [5]. The US national health care spending associated with hypertension was estimated to be approximately US $131 billion, averaged over 12 years from 2003 to 2014 [6]. In 2017, the total estimated cost of diagnosed diabetes in the United States was US $327 billion, including medical costs and lost productivity, and provision of care for people with diagnosed diabetes accounted for a quarter of health care costs [7]. Without appropriate management of obesity, hypertension, and T2DM, patients experience disastrous complications, and societies are troubled with HCEs and disease-related productivity losses.

Self-management is crucial for the proper management of the 3 chronic diseases. A structured lifestyle intervention program comprising a healthy diet, physical activity, and behavioral interventions is essential for the treatment of obesity [8]. Lifestyle management and self-management with self-monitoring are also important in the treatment of hypertension and T2DM [9-11]. However, in face-to-face outpatient consultations, health care providers lack time to deliver information and skills for self-management to patients and motivate them to change their lifestyles.

In this context, mobile health (mHealth) can be a cost-effective tool for self-management in the treatment of chronic diseases. The Global Observatory for eHealth defined mHealth as a medical and public health practice supported by mobile devices [12]. It is useful as it can (1) enhance drug adherence through reminders, (2) facilitate self-monitoring coupled with wireless medical peripheral devices, and (3) provide tailored practical information.

A considerable number of clinical trials have been executed to inspect the usefulness of mHealth interventions in the treatment of obesity, hypertension, and T2DM [13,14]. The reviews and meta-analyses of these trials indicate that although mHealth interventions are likely to promote weight loss, lower hemoglobin A1c (HbA_{1c}), and reduce blood pressure, the individual results are mixed [13,14]. In Korea, several groups have reported the clinical usefulness of smartphone-based apps in T2DM patients; however, their glucose-lowering effect is not clear [15-17]. Although numerous trials have investigated the efficacy of mHealth in the management of obesity, hypertension, and T2DM, they have explored its effect with respect to the 3 diseases, separately or in patients with obesity and hypertension or in patients with hypertension and T2DM. As hypertension and T2DM are comorbidities of obesity, and the 3 diseases are important risk factors of coronary vascular disease, integrative lifestyle approaches for the 3 diseases are more appropriate. They should include the feedback system of diet and exercise, medication assistance, and self-monitoring of blood pressure, blood glucose, and body composition.

Objectives

The aim of this 6-month crossover pilot study is to evaluate the clinical effects of integrative mHealth supported by self-monitoring home devices among patients with T2DM or hypertension and obesity. The integrative mHealth used in this pilot study provides a platform to link the out-of-hospital self-monitoring results of diet, exercise, blood pressure, blood glucose, and body composition with web servers for data storage and web portals for the patient and their physician's data access. Embedded apps in patients' smartphones are LIBIT (Huraypositive Co) for recording diet and exercise, connecting self-measuring home devices to web servers, and providing feedback and health information to patients; and Mediram (GST Korea) for medication assistance. These smartphone apps have been newly developed for this project.

Methods

Study Participants

The pilot study was conducted with adults aged 40 to 70 years who were treated for T2DM (without insulin) or hypertension in the departments of family medicine and endocrinology at 2 university hospitals and were in stable status for at least the past 4 months. Recent HbA_{1c} levels of participants measured in <4 months were >6.0%. To use the Bluetooth-enabled self-measuring home devices and smartphone apps developed for this pilot study, the participants should have had and been able to use Android smartphones with OS version 4.3 (jellybean) or later. Recruitment was conducted between October 2018 and February 2020 in Incheon and Daejeon, which are 2 metropolitan cities in South Korea. The exclusion criteria comprised a history of malignant diseases, coronary artery obstructive disease, stroke, organ transplantation, drug abuse and alcohol dependence, disability or respiratory disease limiting

XSL•FO
RenderX

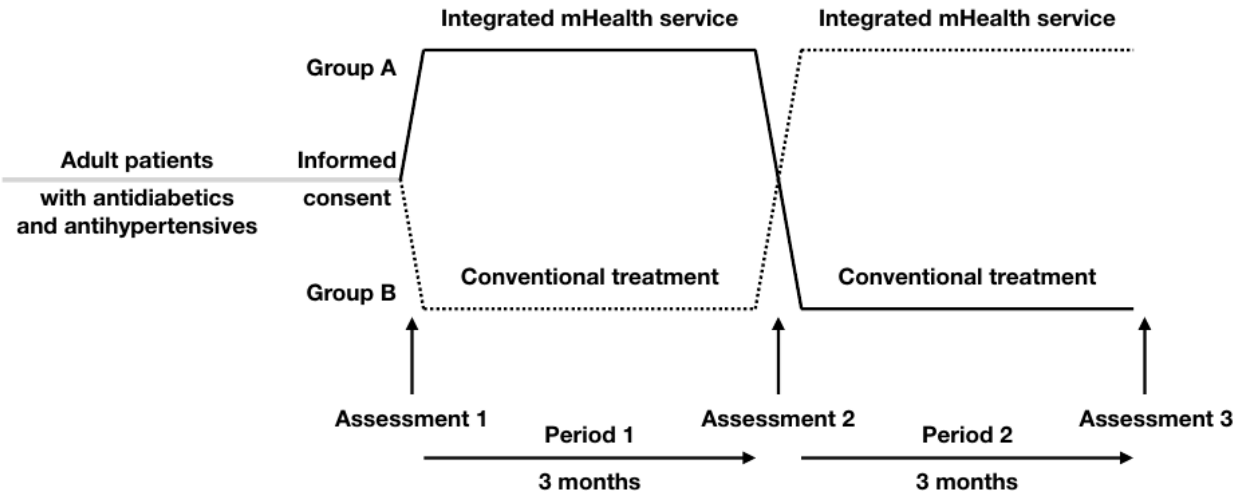
exercise, and hospitalization in the past 6 months with major medical conditions.

All participants were informed of the aim and process of the pilot study during the interviews and were requested for their consent to join this study. The study procedures were performed only with participants who provided informed consent. This pilot study was approved by the institutional review boards of Gachon University Gil Medical Center and Chungnam National University Hospital, and the pilot study was performed in accordance with the Declaration of Helsinki and the guidelines of Good Clinical Practice. Deidentified and anonymized data were used in the analyses.

Study Design, Devices, and Smartphone Apps

This pilot study was performed using a controlled, randomized, 3-month, 2-period crossover design to test the efficacy of integrative mHealth coupled with self-monitoring home devices and smartphone apps as opposed to the conventional treatment (CON) of T2DM with or without hypertension. Figure 1 demonstrates the pilot study design. The recruited participants were randomly assigned to 2 groups using computer-generated random numbers: 1 group started with the integrative mHealth service period and switched to the CON period (mHealth-CON group), whereas the other group started with the CON period and switched to the mHealth period (CON-mHealth group). There was no washout period. Measurements of body weight, body composition, blood pressure, and HbA_{1c} were taken at the start of the first treatment period, during treatment transition, and at the end of the second treatment period. A survey on smartphone apps was conducted at the end of the pilot study.

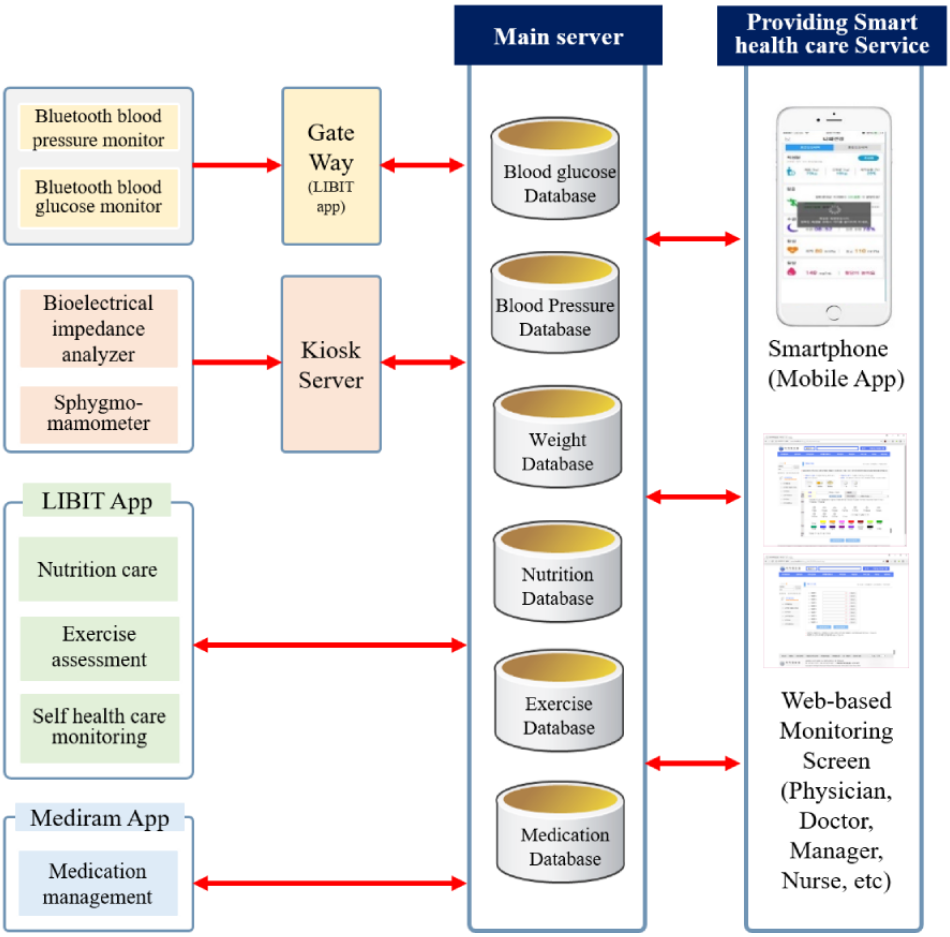
Figure 1. Study design. mHealth: mobile health.



Integrative mHealth supported the participants and physicians through the following 4 components: (1) self-measuring devices; (2) smartphone apps that gathered and transferred data on the participant's lifestyle and provided feedback, health and drug information, and reminders of their medication schedule; (3) unmanned kiosks for the official measurement of blood pressure and body composition; and (4) web-based access to participants' health information through which physicians could review participants' health data at a glance. The entire architecture of the information transmission in this pilot study is shown in Figure 2. For systematic collection and administration of health information data, this pilot study emphasized data security by applying 5 systems: (1) section encoding via secure socket layer (SSL) or transport layer security, (2) encoding critical information, (3) controlling the users' and administrators' accessibility to data, (4) restricting the collection of personal identification information, and (5) agreeing to collect and use personal identification information. Information with a high risk of data loss, such as passwords, was saved using the unilateral encoding system of Secure Hash Algorithm 256 in the health information service system. Integrative mHealth was available only for the mHealth period. The devices and smartphone apps were supplied at the commencement of the mHealth period and retrieved at the end of it.

XSL•FO
RenderX

Figure 2. Architecture of information transmission.



A total of 2 Bluetooth-enabled devices, the blood pressure monitor HEM-9200T (Omron) and the blood glucose monitor CareSense N Premier BLE (i-sense), were used as self-measuring home devices in this pilot study (Multimedia Appendix 1). A total of 2 unmanned digital kiosks (GST Korea) were established at both hospitals for the self-measurement of blood pressure and body composition (Multimedia Appendix 2). The kiosks were equipped with a user screen, a radio-frequency identification card reader, a bioelectric impedance analyzer SC-330 (Tanita), and an automatic upper arm sphygmomanometer BP-210 (Accuniq). For using the kiosks, a radio-frequency identification card was supplied to each participant.

The readings of blood pressure and glucose levels that were measured at home were transmitted to the patient's Android smartphone app LIBIT through a Bluetooth connection and then transferred to the main server using cellular data. The data on body composition and blood pressure measured at the kiosks were also linked to the main server through http secure based on certificate verification; http secure sent encoded information of the clients to the server using the security protocol of SSL. SSL or transport layer security operates in the same way as a virtual private network, sending security data to the server via a virtual tunnel.

A total of 2 Android apps, LIBIT and Mediram, were developed to gather participants' lifestyle data, provide feedback and health and drug information, and enhance their adherence to medication in this pilot study. The LIBIT app comprised 4 functions: nutrition care, exercise assessment, transmission of self-monitored blood pressure and glucose level, and health monitoring (Multimedia Appendix 3). The nutrition care component of LIBIT calculated the suggested total calorie and macronutrient ratios of each participant. Users could record their food intake through the smartphone keypad or their voice using embedded voice recognition technology. Analyzing the food intake records, LIBIT estimated the intake of 14 nutrients (total calories, carbohydrates, proteins, fats, calcium, phosphorus, iron, potassium, sodium, vitamin A, thiamine, riboflavin, niacin, and vitamin C) and reported each of them as insufficient, suitable, or excessive for the users. Users could also record the kind and duration of exercise through the exercise assessment function of LIBIT; it subsequently calculated the amount of calorie consumption and reported it to the users. LIBIT received participants' self-monitored data on blood pressure and glucose from the peripheral devices through a Bluetooth connection and transmitted them to the main server using cellular data. The health monitoring function of LIBIT provided visual feedback of users' health status by creating trend graphs of body weight, blood pressure, glucose level, and calorie consumption. The Mediram app offered comprehensive medication information to users (Multimedia Appendix 4). Users could easily upload the prescription to Mediram by just scanning the QR codes of their prescriptions. Mediram supplied drug information and notified users of their drug schedule to enhance their adherence to medication.

XSL•FO
RenderX

Physicians could access the participants' health data using Bluetooth-enabled home devices and unmanned kiosks at the main server using web browsers (Multimedia Appendix 5). The average blood pressure, glucose level, adherence to medication, body composition change, nutrient intake, and calorie consumption through exercise in the given period were displayed on a page, and the physician could monitor changes in a participant's health status at a glance.

Statistical Analyses

The input rate of food intake was calculated as the total input frequency divided by the product of the days of the mHealth period and frequency of daily food intake. The input rate of taking medicine was calculated similarly as the total input frequency divided by the product of days of the mHealth period and frequency of daily drug intake. The input rate of exercise was calculated as the total input frequency divided by the number of days of the mHealth period.

Student *t* test, chi-square test, and Fisher exact test were used to determine the differences in baseline characteristics between the mHealth-CON and CON-mHealth groups. The Wilcoxon signed-rank test was conducted to test the difference in the input rates of diet, exercise, and medicine intake between the groups. A paired *t* test or Wilcoxon signed-rank test was performed to determine the differences in changes between the integrative mHealth and CON periods. Subsequently, according to the input rate of medicine intake, the estimated between-group differences in changes in the variables—obesity, hypertension, and T2DM—were calculated using generalized linear models, after adjusting for treatment (mHealth or CON) in model 1; treatment, group (mHealth-CON or CON-mHealth), and sex in model 2; and treatment, group, sex, and age in model 3. To compare the 3 input functions of the apps, a Kruskal–Wallis rank sum test [18] was conducted.

All statistical analyses were implemented in the R software version 4.0.3 (R Core Team), which is a language and environment for statistical computing. A 2-tailed $P<.05$ was considered statistically significant.

Results

Baseline Characteristics

A total of 37 participants were enrolled in this pilot study. Among the 37 participants, there were 3 (8%) cases of newly diagnosed T2DM, 1 (3%) case of bioelectrical impedance analysis showing error, and 1 (3%) case of dropout. In the final data set, 32 participants' data were included, of whom 15 (47%) were allocated to the mHealth-CON group, whereas the remaining 17 (53%) were assigned to the CON-mHealth group.

Among the 32 participants, 23 (72%) were men. The mean age of participants was 56.8 years, ranging from 40 to 69 years. Of the 32 participants, 17 (53%) graduated from college or above. Approximately 38% (12/32) of participants had hypertension. Most participants were overweight and obese. According to the BMI classification of the Korean Society for the Study of Obesity [19], only 3% (1/32) of participants were in the normal range, that is, 20.8 kg/m^2. Of the 32 participants, 7 (22%) were overweight, with BMIs ranging from 23 kg/m^2 to 24.9 kg/m^2; 20 (63%) had class 1 obesity, with a BMI range of 25 kg/m^2 to 29.9 kg/m^2; and BMIs of the remaining 4 (13%) participants were >30 kg/m^2. The mean HbA_{1c} level was 7.5%, ranging from 6.1% to 9.4%. The baseline demographic and clinical characteristics of the 2 groups, mHealth-CON and CON-mHealth, were similar, except for the frequency of hypertension (Table 1). There were no significant differences in demographic characteristics, body weight and body composition, blood pressure, and HbA_{1c} levels between the groups.

XSL•FO
RenderX

Table 1. Participant baseline characteristics (N=32).

Group	mHealth[a]-CON[b] (n=15)	CON-mHealth (n=17)	*P* value[c]
Sex, n (%)			
Male	9 (60)	14 (82)	.24
Age (years), mean (SD)	58.9 (4.9)	55.1 (7.6)	.11
Education, n (%)			.14
Elementary school	1 (7)	1 (6)	
Middle school	0 (0)	2 (12)	
High school	8 (53)	3 (18)	
College or above	6 (40)	11 (65)	
Marital status, n (%)			.99
Divorced	1 (7)	1 (6)	
Married	14 (93)	16 (94)	
House, n (%)			.40
Apartment	9 (60)	10 (63)	
Detached	2 (13)	5 (31)	
Unit	2 (13)	1 (6)	
Other	2 (13)	0 (0)	
Monthly income[d], mean (SD)	5.3 (2.3)	5.6 (2.3)	.71
Hypertension, mean (SD)	9 (60)	3 (18)	.01[e]
Body weight (kg), mean (SD)	73.4 (9.4)	80.0 (11.5)	.09
BMI (kg/m^2), mean (SD)	26.8 (2.3)	27.8 (3.6)	.34
Fat mass (kg), mean (SD)	22.3 (5.5)	22.8 (6.9)	.83
Body fat (%), mean (SD)	30.5 (7.1)	28.4 (6.9)	.38
Fat free mass (kg), mean (SD)	51.1 (9.0)	57.2 (9.5)	.07
Body water (%), mean (SD)	51.8 (4.1)	53.0 (4.8)	.45
Blood pressure (mm Hg), mean (SD)			
Systolic	134.8 (11.4)	129.4 (13.2)	.23
Diastolic	81.5 (7.8)	78.8 (11.8)	.46
HbA_{1c}[f] (%), mean (SD)	7.5 (0.7)	7.5 (0.8)	.98

[a]mHealth: mobile health.

[b]CON: conventional treatment.

[c]Calculated using Fisher exact test or Student *t* test.

[d]1: none, 2: <1 million Korean won (KRW), 3: KRW 1-2 million, 4: KRW 2-3 million, 5: KRW 3-4 million, 6: KRW 4-5 million, 7: KRW 5-6 million, 8: > KRW 6 million.

[e]Calculated using chi-square test.

[f]HbA_{1c}: hemoglobin A_{1c}.

Diet and Exercise

The input rates of food intake (24.9%) and exercise (5.3%) were very low (Table 2). Moreover, both input rates were significantly low in the CON-mHealth group, which means that there was attrition in food intake and exercise input to the LIBIT app over time during the pilot study. On account of low input rates of food intake and exercise, it was impossible to execute the analysis of data related to nutrient intake and energy consumption.

XSL•FO
RenderX

Table 2. Input rates of diet, exercise, and taking medicine (N=32).

Parameters	mHealth[a]-CON[b]	CON-mHealth	Total	*P* value[c]
Diet, mean (SD)	36.6 (39.3)	14.5 (23.7)	24.9 (33.3)	.03
Exercise, mean (SD)	8.0 (9.3)	2.9 (8.9)	5.3 (9.3)	.002
Taking medicine, mean (SD)	80.3 (20.0)	87.3 (20.8)	84.0 (20.4)	.06

[a]mHealth: mobile health.
[b]CON: conventional treatment.
[c]Calculated using Wilcoxon rank sum test.

Drug Adherence

The input rate of medicine intake to the Mediram app was high, at 84.0%. Unlike the input rates of food intake and exercise, there was no attrition in the input of medicine intake. There was no significant difference in the input rate of taking medicine irrespective of whether the mHealth period was before or after the CON period (80.3% and 87.3%, respectively; *P*=.06; Table 2).

Efficacy of mHealth on Body Weight, Body Composition, Blood Pressure, and HbA_{1c}

The changes in body weight and body composition, blood pressure, and HbA_{1c} between the mHealth and conventional periods are displayed in Table 3. There were no significant differences in the changes in the variables of obesity, hypertension, and T2DM. The individual changes in these variables are shown in Multimedia Appendix 6.

Table 3. Comparison of the changes in body weight, body composition, blood pressure, and HbA_{1c}[a].

Characteristics	N (mHealth[b]/CON[c])	mHealth, mean (SD)	Conventional, mean (SD)	*P* value[d]
Body weight (kg)	32/32	–0.519 (1.655)	0.000 (1.832)	.29
BMI (kg/m^2)	32/32	–0.133 (0.640)	–0.167 (0.709)	.86
Fat mass (kg)	31/30	–0.292 (1.964)	0.153 (2.357)	.56[e]
Body fat (%)	31/30	–0.255 (2.454)	0.172 (2.574)	.43[e]
Fat free mass (kg)	31/30	–0.102 (1.703)	–0.173 (1.933)	.78[e]
Body water (%)	30/30	0.240 (3.173)	–0.360 (2.702)	.21
Blood pressure (mm Hg)				
Systolic	31/31	–0.226 (12.328)	–2.839 (11.097)	.51
Diastolic	31/31	–2.839 (9.000)	–1.097 (9.239)	.51[e]
HbA_{1c}[f] (%)	32/32	–0.269 (0.663)	–0.009 (0.693)	.19

[a]Some participants' data are missing.
[b]mHealth: integrative mobile health service.
[c]CON: conventional treatment.
[d]Calculated by paired *t* test.
[e]Calculated using Wilcoxon signed-rank test.
[f]HbA_{1c}: hemoglobin A_{1c}.

Effect of the Medication Assistance App on Body Weight, Body Composition, Blood Pressure, and HbA_{1c}

To inspect the effect of mHealth apps on body weight and composition, blood pressure, and HbA_{1c}, the estimated changes in variables, according to the input rate of medicine intake, were calculated using generalized linear models (Table 4). *Group* was included as a variable in models 2 and 3 as there were differences in the prevalence of hypertension between the groups mHealth-CON and CON-mHealth. In proportion to the elevation of the input rate of medicine intake, body fat mass and HbA_{1c} were lower (Table 4). Owing to low input rates, the effects of food intake and exercise-related app functions on changes in the variables could not be assessed.

XSL•FO
RenderX

Table 4. The estimated changes in variables—obesity, hypertension, and type 2 diabetes mellitus—according to the input rate of taking medicine.

Characteristics	Model 1[a]		Model 2[b]		Model 3[c]	
	Estimate	*P* value	Estimate	*P* value	Estimate	*P* value
Body weight (kg)	–0.016	.16	–0.019	.09	–0.019	.11
BMI (kg/m^2)	–0.004	.34	–0.005	.25	–0.005	.31
Fat mass (kg)	–0.027	.08	–0.034	.03	–0.032	.04
Body fat (%)	–0.030	.10	–0.037	.05	–0.035	.07
Fat free mass (kg)	0.007	.61	0.010	.49	0.008	.55
Body water (%)	0.035	.11	0.042	.07	0.040	.08
Blood pressure (mm Hg)						
Systolic	–0.006	.95	–0.016	.85	–0.014	.87
Diastolic	–0.016	.80	–0.028	.67	–0.028	.68
HbA_{1c}[d] (%)	–0.008	.07	–0.009	.04	–0.010	.03

[a]Model 1 adjusted for treatment (mobile health [mHealth] or conventional treatment [CON]).

[b]Model 2 adjusted for treatment, group (mHealth-CON or CON-mHealth), and sex.

[c]Model 3 adjusted for treatment, group, sex, and age.

[d]HbA_{1c}: hemoglobin A_{1c}.

Survey on the Smartphone Apps

The results of the smartphone app survey at the end of the pilot study are summarized in Tables 5 and 6. Among the 3 input functions of food intake, exercise, and medicine intake in smartphone apps, the input of medicine intake was a more helpful, easier to use, and better-designed function than the others. There were more opinions about improvements in the input of food intake. The 2 most difficult functions were those of recording food intake and finding food items in the provided list. For the input of exercise, multitasking with other apps was highly desired.

Table 5. Survey about the functions of smartphone apps for the input of food intake, exercise, and taking medicine (Likert scale result; N=32).

Function of the apps and question	Likert scale (1-5), mean (SD)
Input of food intake	
Helpful[a]	3.4 (1.0)
Easy[b]	2.9 (1.2)
Well-functioned[c]	3.1 (0.9)
Input of exercise	
Helpful[a]	3.6 (1.0)
Easy[b]	3.3 (1.0)
Well-functioned[c]	3.3 (0.9)
Input of taking medicine	
Helpful[a]	4.1 (0.8)
Easy[b]	4.1 (0.7)
Well-functioned[c]	4.0 (0.9)

[a]P=.01 using the Kruskal-Wallis rank sum test.

[b]P<.001 using the Kruskal-Wallis rank sum test.

[c]P=.002 using the Kruskal-Wallis rank sum test.

XSL•FO
RenderX

Table 6. Survey about the functions of smartphone apps for the input of food intake, exercise, and taking medicine (opinions; N=32).

Opinions	Numbers, n (%)
Input of food intake	
Have experience of using the diet app	1 (3)
What needs to be improved?	12 (38)
Difficult to record food intake	4 (13)
Errors in voice recognition	1 (3)
Wish the portion size was broken down	2 (6)
Wish to input data on my own	1 (3)
No food in the food list	5 (16)
Unbelievable calculated calorie	1 (3)
Input of exercise	
What needs to be improved?	7 (22)
Wish it worked with other apps simultaneously	4 (13)
Wish to input data on my own	1 (3)
Wish it was recorded automatically	2 (6)
Difficult to input data	1 (3)
Input of taking medicine	
What needs to be improved?	5 (16)
Errors in running the app	1 (3)
Not easy to use	2 (6)
Wish to go back to home screen after the input	1 (3)
Wish to control the medication time	1 (3)

Discussion

Principal Findings

This pilot study demonstrated that smartphone apps could influence changes in body fat and blood glucose levels. As the input rate of medicine intake increased, body fat mass and HbA_{1c} decreased. Although the improvement in drug adherence for diabetes is expected to enhance the control of blood glucose levels, good adherence to antidiabetes medication is irrelevant to body fat reduction; however, some—not all—antidiabetes medicines can induce body weight loss [20]. This relationship can be better explained by positive behavioral changes and effective self-management skills obtained from the integrative mHealth intervention. These findings imply that mHealth can improve body fat and blood glucose status in patients with T2DM or hypertension; however, it failed to result in clinical improvement in this pilot study. In addition, an app for drug information and reminders is more pleasing to the eyes of the patients than a diet diary and exercise monitor. A larger and long-term clinical trial is needed to determine whether integrative mHealth services help patients with T2DM or hypertension and obesity. The findings of this pilot study are currently being applied to an improved mHealth intervention project for people who have moved into a large, new apartment complex.

A special feature of this pilot study is its crossover design. This pilot study allocated participants to 2 different interventions over two 3-month periods. A crossover design has the following advantages over a parallel design: (1) it may offer more precise estimates of intervention effects as it would remove the differences in participants' characteristics and methodological variations of open trials, and (2) it requires a smaller number of participants [21,22]. Obesity, hypertension, and T2DM are chronic diseases and are appropriate for a crossover study as the conditions of patients are stable if the prescriptions do not change, and they are not usually curable. In addition, a crossover design was appropriate as integrative mHealth service was an add-on treatment to the CON and not a separate stand-alone treatment.

Although there was a chance of carryover effect in this pilot study, it could not have widened the differences in changes between the treatment groups. A washout period was logically impossible for the transition from CON to the add-on integrative mHealth intervention. There might be a carryover effect in the mHealth-CON group, which could be the reason that integrative mHealth could not induce significant clinical improvement as compared with CON alone. According to previous studies [14,23], mHealth has a clinically positive effect on chronic disease management. In the case of the carryover effect, it would strengthen the positive effect of the second phase of the treatment, that is, CON; subsequently, the differences in the 2

XSL•FO
RenderX

treatment phases would become smaller, which would make the analysis more conservative.

The input rate of taking drugs was high, that is, >80%, even in the CON-mHealth group, where Mediram use was assigned in the second 3-month period. In fact, it was marginally higher in the CON-mHealth group than in the mHealth-CON group. Pharmacological adherence is very important for the successful management of hypertension and diabetes. Among adults with several common chronic diseases, only 40% to 70% of medications are taken properly [24,25]. Poor drug adherence might lead to poor clinical outcomes and increased HCE in chronic diseases [26]. The use of mHealth may increase medication adherence in chronic diseases and coronary heart disease; however, the results are diverse [27-29]. The variation among the previous trials was probably because of the different modules of mHealth that facilitated proper drug intake. It appears that the medication assistance app Mediram independently improved body fat and blood glucose status in our pilot study. Further research is needed to verify the clinical usefulness of the medication assistance app and the features of the app that make it more effective.

On the contrary, the nutrition and exercise care app LIBIT was not popular among most participants. Although LIBIT introduced a voice recognition technique for diet diaries to make food intake input much easier, the mean input rate of food intake was only 24.9%. The self-recording rate of exercise was even lower, at 5.3%. Following the low input rates of food intake and exercise, individualized lifestyle feedback based on the amount of macro- and micronutrient intake and calorie consumption through exercise was not properly reported to the participants. Most of all, for the individualized lifestyle feedback system, the input methods of diet and activity should be easy and simple, as revealed by the participants.

The integrative mHealth service of this pilot study failed to clinically reduce body weight and fat, lower blood pressure, and improve T2DM. The amount of HbA_{1c} reduction in the mHealth period in our pilot study (–0.269%) was smaller than that in previous reports. A total of 2 meta-analyses reported that mHealth interventions improved HbA_{1c} significantly, with mean differences of –0.39% and –0.44%, respectively [14,23]. However, for obesity and blood pressure, the results of previous studies were mixed, and there might be attrition of improvement over time. A randomized controlled trial for 3 months in patients with T2DM and hypertension reported that the blood pressure differences between the intervention and control groups were narrowed during the second and third months compared with that of the first month [30]. Wang et al [13] outlined that mHealth induced an average weight loss widely ranging from –1.97 kg in 16 weeks to –7.1 kg in 5 weeks and mentioned that most studies were executed with small samples and in short intervention periods and did not use proper data collection or analytical methods. A reason for the failure of our mHealth intervention in improving body fat, blood pressure, and blood glucose levels might be the unsuccessful individualized lifestyle modification linked with the neglected input of food intake and exercise, considering the high input rate of medicine intake linked with the decrease in body fat and HbA_{1c}.

Diabetes and hypertension are 2 major chronic diseases that incur burdens on public health economically, and the burdens will be bigger in the near future of the aging world. Life expectancy (LE) has increased drastically over the past several decades. Between 1950 and 2017, it increased from 48.1 years to 70.5 years for men and from 52.9 years to 75.6 years for women worldwide [31]. However, there is a big gap between LE and health-adjusted LE (HALE). Although the global average HALE increased from 57.6 years in 1995 to 63.3 years in 2017, the gap between LE and HALE also increased from 8.6 years to 9.7 years during the same period [32]. Population aging has been considered a big source of increase in HCE, and individual health status has been suggested as a main factor of HCE in the aging population [33]. Older adults who are hypertensive patients are more likely to have complications, including congestive heart failure or chronic kidney disease, and these comorbidities induce incremental medical expenditures for adults aged ≥65 years, which is approximately US $2500 more than that for adults aged 18 to 44 years [6]. Similarly, the prevalence of diabetes and the medical costs related to diabetes are primarily increasing among the population aged ≥65 years [7]. Integrative mHealth can be a cost-effective tool to prevent catastrophic complications and increased HCE, which are associated with hypertension and diabetes management in the aging population, by enhancing their self-monitoring skills and drug adherence.

There were some limitations to our pilot study. First, the treatment period was short, at 3 months. Until now, the long-term efficacy of mHealth has been doubtful, and one cannot be sure if its effect would wear down over time. Long-term clinical trials with serial assessments of their effects are necessary for the future. Second, only a few participants were prompt in recording their input of food intake and exercise on the lifestyle app. Consequently, without the availability of individualized lifestyle information, the integrative mHealth service in this pilot study was scaled down to the combination of self-monitoring of blood pressure and glucose levels, medication assistance app, unmanned kiosks, and physicians' access to participants' health records through a web browser. In addition to simpler and easier input methods of diet and exercise, more immersive and highly functional mHealth apps should be designed that focus in depth on the content and user experience and can motivate patients. Third, bioelectrical impedance analysis was used for the measurement of body composition. For the assessment of treatment effects on body composition, dual-energy X-ray absorptiometry would be appropriate, as bioelectrical impedance analysis is easily affected by the hydration status of the body. Finally, this pilot study was conducted with relatively young adults and not with older adults. Only 6% (2/32) of participants were aged >65 years. As chronic disorders are more common among older adults who may not have good digital literacy and have difficulty in adopting new information technologies, the apps and peripheral medical devices that are designed for self-monitoring should be better accessible for older adults. In addition, voice-based mHealth may be preferred by older adults with limited digital literacy and poor vision [34].

XSL•FO
RenderX

Conclusions

This pilot study illustrated that smartphone apps could influence changes in body fat and blood glucose; however, mHealth failed to result in clinical improvement. A higher input rate of medicine intake was related to a significantly lower body fat mass and HbA_{1c}. This result could possibly be because of positive behavioral changes and effective self-management skills obtained from the integrative mHealth intervention. In addition, the app for drug information and reminders was considered to be more pleasing to the eyes of the patients than a diet diary and exercise monitor.

Acknowledgments

The authors appreciate Huraypositive Co for providing the LIBIT app and GST Korea for supplying the Mediram app and nonhuman kiosks.

This work was supported by a grant (19RERP-B090094-06) from the Residential Environment Research Program funded by the Ministry of Land, Infrastructure, and Transport of the Korean government (NRF-2017R1A6A3A11036413) and a grant from the National Research Foundation of Korea funded by the Korean government.

Authors' Contributions

SWO conceptualized, funded, and administered the pilot study; KKK and SSK executed the clinical trial; KKK and SKP acquired and analyzed the data; KKK and SP interpreted data; KKK drafted the manuscript; and KKK, SSK, SP, and SWO revised the manuscript with critics.

Conflicts of Interest

None declared.

Multimedia Appendix 1
(a) Bluetooth blood pressure monitor and (b) Bluetooth blood glucose monitor.
[PNG File , 495 KB - mhealth_v10i1e27192_app1.png]

Multimedia Appendix 2
Nonhuman kiosks comprising a bioelectrical impedance analyzer, a user screen, a mobile health service radio-frequency identification card reader, and an automatic upper arm blood pressure monitor.
[PNG File , 383 KB - mhealth_v10i1e27192_app2.png]

Multimedia Appendix 3
Smartphone screen of the LIBIT app for (a) nutrition and (b) exercise assessment.
[PNG File , 485 KB - mhealth_v10i1e27192_app3.png]

Multimedia Appendix 4
Smartphone screen of the Mediram app.
[PNG File , 330 KB - mhealth_v10i1e27192_app4.png]

Multimedia Appendix 5
Participant’s health data access through web browser.
[PNG File , 263 KB - mhealth_v10i1e27192_app5.png]

Multimedia Appendix 6
Line plots of individual changes in (a) body weight, (b) BMI, (c) body fat mass, (d) body fat percentage, (e) fat free mass, (f) body water percentage, (g) systolic blood pressure, (h) diastolic blood pressure, and (i) HbA_{1c}. mHealth: integrative mobile health service; CON: conventional treatment.
[PNG File , 2117 KB - mhealth_v10i1e27192_app6.png]

References

1. Kotsis V, Jordan J, Micic D, Finer N, Leitner DR, Toplak H, et al. Obesity and cardiovascular risk: a call for action from the European Society of Hypertension Working Group of Obesity, Diabetes and the High-risk Patient and European Association for the Study of Obesity: part A: mechanisms of obesity induced hypertension, diabetes and dyslipidemia and

XSL•FO
RenderX

practice guidelines for treatment. J Hypertens 2018 Jul;36(7):1427-1440. [doi: 10.1097/HJH.0000000000001730] [Medline: 29634663]
2. Wilson PW. Established risk factors and coronary artery disease: the Framingham Study. Am J Hypertens 1994 Jul;7(7 Pt 2):7-12. [doi: 10.1093/ajh/7.7.7s] [Medline: 7946184]
3. Sowers JR. Obesity as a cardiovascular risk factor. Am J Med 2003 Dec 08;115 Suppl 8A(8):37-41. [doi: 10.1016/j.amjmed.2003.08.012] [Medline: 14678864]
4. Withrow D, Alter D. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes Rev 2011 Feb;12(2):131-141. [doi: 10.1111/j.1467-789X.2009.00712.x] [Medline: 20122135]
5. Song HJ, Hwang J, Pi S, Ahn S, Heo Y, Park S, et al. The impact of obesity and overweight on medical expenditures and disease incidence in Korea from 2002 to 2013. PLoS One 2018;13(5):e0197057 [FREE Full text] [doi: 10.1371/journal.pone.0197057] [Medline: 29746509]
6. Kirkland EB, Heincelman M, Bishu KG, Schumann SO, Schreiner A, Axon RN, et al. Trends in healthcare expenditures among US adults with hypertension: national estimates, 2003–2014. J Am Heart Assoc 2018 Jun 05;7(11):e008731. [doi: 10.1161/jaha.118.008731]
7. American Diabetes Association. Economic costs of diabetes in the U.S. in 2017. Diabetes Care 2018 Dec;41(5):917-928. [doi: 10.2337/dci18-0007] [Medline: 29567642]
8. Garvey WT, Mechanick JI, Brett EM, Garber AJ, Hurley DL, Jastreboff AM, Reviewers of the AACE/ACE Obesity Clinical Practice Guidelines. American Association of Clinical Endocrinologists and American College of Endocrinology Comprehensive Clinical Practice Guidelines for medical care of patients with obesity. Endocr Pract 2016 Jul;22 Suppl 3:1-203. [doi: 10.4158/EP161365.GL] [Medline: 27219496]
9. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Houston MN, Hubbard VS, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014 Jul 01;63(25 Pt B):2960-2984 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.003] [Medline: 24239922]
10. Shahaj O, Denneny D, Schwappach A, Pearce G, Epiphaniou E, Parke HL, et al. Supporting self-management for people with hypertension: a meta-review of quantitative and qualitative systematic reviews. J Hypertens 2019 Feb;37(2):264-279. [doi: 10.1097/HJH.0000000000001867] [Medline: 30020240]
11. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. Management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 2018 Dec;41(12):2669-2701. [doi: 10.2337/dci18-0033] [Medline: 30291106]
12. mHealth: new horizons for health through mobile technologies. World Health Organization. 2011. URL: https://www.afro.who.int/publications/mhealth-new-horizons-health-through-mobile-technologie [accessed 2021-12-11]
13. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. Adv Nutr 2017 May;8(3):449-462. [doi: 10.3945/an.116.014100] [Medline: 28507010]
14. Liu K, Xie Z, Or CK. Effectiveness of mobile app-assisted self-care interventions for improving patient outcomes in type 2 diabetes and/or hypertension: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Aug 04;8(8):e15779 [FREE Full text] [doi: 10.2196/15779] [Medline: 32459654]
15. Kim H, Choi W, Baek EK, Kim YA, Yang SJ, Choi IY, et al. Efficacy of the smartphone-based glucose management application stratified by user satisfaction. Diabetes Metab J 2014 Jun;38(3):204-210 [FREE Full text] [doi: 10.4093/dmj.2014.38.3.204] [Medline: 25003074]
16. Kim YJ, Rhee SY, Byun JK, Park SY, Hong SM, Chin SO, et al. A smartphone application significantly improved diabetes self-care activities with high user satisfaction. Diabetes Metab J 2015 Jun;39(3):207-217 [FREE Full text] [doi: 10.4093/dmj.2015.39.3.207] [Medline: 26124991]
17. Yang Y, Lee EY, Kim H, Lee S, Yoon K, Cho J. Effect of a mobile phone-based glucose-monitoring and feedback system for type 2 diabetes management in multiple primary care clinic settings: cluster randomized controlled trial. JMIR Mhealth Uhealth 2020 Feb 26;8(2):e16266 [FREE Full text] [doi: 10.2196/16266] [Medline: 32130172]
18. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952 Dec;47(260):583-621. [doi: 10.1080/01621459.1952.10483441]
19. Seo MH, Lee W, Kim SS, Kang J, Kang J, Kim KK, Committee of Clinical Practice Guidelines, Korean Society for the Study of Obesity (KSSO). 2018 Korean society for the study of obesity guideline for the management of obesity in Korea. J Obes Metab Syndr 2019 Mar;28(1):40-45 [FREE Full text] [doi: 10.7570/jomes.2019.28.1.40] [Medline: 31089578]
20. Lau DC, Teoh H. Impact of current and emerging glucose-lowering drugs on body weight in type 2 diabetes. Can J Diabetes 2015 Dec;39 Suppl 5:148-154. [doi: 10.1016/j.jcjd.2015.09.090] [Medline: 26654858]
21. Richens A. Proof of efficacy trials: cross-over versus parallel-group. Epilep Res 2001 May 1;45(1-3):43-47 [FREE Full text] [doi: 10.1016/s0920-1211(01)00214-5]
22. Mills EJ, Chan A, Wu P, Vail A, Guyatt GH, Altman DG. Design, analysis, and presentation of crossover trials. Trials 2009 Apr 30;10(1):27 [FREE Full text] [doi: 10.1186/1745-6215-10-27] [Medline: 19405975]

XSL•FO
RenderX

23. Mao Y, Lin W, Wen J, Chen G. Impact and efficacy of mobile health intervention in the management of diabetes and hypertension: a systematic review and meta-analysis. BMJ Open Diabetes Res Care 2020 Sep;8(1) [FREE Full text] [doi: 10.1136/bmjdrc-2020-001225] [Medline: 32988849]
24. Briesacher B, Andrade S, Fouayzi H, Chan K. Comparison of drug adherence rates among patients with seven different medical conditions. Pharmacotherapy 2008 Apr;28(4):437-443 [FREE Full text] [doi: 10.1592/phco.28.4.437] [Medline: 18363527]
25. Naderi S, Bestwick J, Wald D. Adherence to drugs that prevent cardiovascular disease: meta-analysis on 376,162 patients. Am J Med 2012 Sep 1;125(9):882-887 [FREE Full text] [doi: 10.1016/j.amjmed.2011.12.013] [Medline: 22748400]
26. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [doi: 10.1097/01.mlr.0000163641.86870.af] [Medline: 15908846]
27. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):52-57 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
28. Mertens A, Brandl C, Miron-Shatz T, Schlick C, Neumann T, Kribben A, et al. A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: a crossover design study comparing documentation via iPad with paper-based control. Medicine (Baltimore) 2016 Sep;95(36):e4446 [FREE Full text] [doi: 10.1097/MD.0000000000004446] [Medline: 27603339]
29. Peng Y, Wang H, Fang Q, Xie L, Shu L, Sun W, et al. Effectiveness of mobile applications on medication adherence in adults with chronic diseases: a systematic review and meta-analysis. J Manag Care Spec Pharm 2020 Apr;26(4):550-561. [doi: 10.18553/jmcp.2020.26.4.550] [Medline: 32223596]
30. Or C, Tao D. A 3-month randomized controlled pilot trial of a patient-centered, computer-based self-monitoring system for the care of type 2 diabetes mellitus and hypertension. J Med Syst 2016 Apr;40(4):81. [doi: 10.1007/s10916-016-0437-1] [Medline: 26802011]
31. Dicker D, Nguyen G, Abate D, Abate KH, Abay SM, Abbafati C, et al. . [doi: 10.1016/s0140-6736(18)31891-9]
32. Cao X, Hou Y, Zhang X, Xu C, Jia P, Sun X, et al. A comparative, correlate analysis and projection of global and regional life expectancy, healthy life expectancy, and their GAP: 1995-2025. J Glob Health 2020 Dec;10(2):020407 [FREE Full text] [doi: 10.7189/jogh.10.020407] [Medline: 33110572]
33. Carreras M, Ibern P, Inoriza JM. Ageing and healthcare expenditures: exploring the role of individual health status. Health Econ 2018 May 09;27(5):865-876. [doi: 10.1002/hec.3635] [Medline: 29424031]
34. Steinman L, Heang H, van Pelt M, Ide N, Cui H, Rao M, et al. Facilitators and barriers to chronic disease self-management and mobile health interventions for people living with diabetes and hypertension in Cambodia: qualitative study. JMIR Mhealth Uhealth 2020 Apr 24;8(4):e13536 [FREE Full text] [doi: 10.2196/13536] [Medline: 32329737]

Abbreviations

CON: conventional treatment
HALE: healthy life expectancy
HbA_{1c}: hemoglobin A_{1c}
HCE: health care expenditure
LE: life expectancy
mHealth: mobile health
SSL: secure socket layer
T2DM: type 2 diabetes mellitus

Edited by L Buis; submitted 15.01.21; peer-reviewed by H Ranjani, K Lakshminarayan; comments to author 19.03.21; revised version received 18.05.21; accepted 19.11.21; published 11.01.22.

Please cite as:
Oh SW, Kim KK, Kim SS, Park SK, Park S
Effect of an Integrative Mobile Health Intervention in Patients With Hypertension and Diabetes: Crossover Study
JMIR Mhealth Uhealth 2022;10(1):e27192
URL: https://mhealth.jmir.org/2022/1/e27192
doi:10.2196/27192
PMID:35014961

©Sang Woo Oh, Kyoung-Kon Kim, Sung Soo Kim, Su Kyung Park, Sangshin Park. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 11.01.2022. This is an open-access article distributed under the terms of the Creative Commons

XSL•FO
RenderX

Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Association Between Home Stay and Symptom Severity in Major Depressive Disorder: Preliminary Findings From a Multicenter Observational Study Using Geolocation Data From Smartphones

Petroula Laiou[1*], PhD; Dzmitry A Kaliukhovich[2*], PhD; Amos A Folarin[1,3,4,5,6], PhD; Yatharth Ranjan[1], MSc; Zulqarnain Rashid[1], PhD; Pauline Conde[1], BSc; Callum Stewart[1], MSc; Shaoxiong Sun[1], PhD; Yuezhou Zhang[1], MSc; Faith Matcham[7], PhD; Alina Ivan[7], MSc; Grace Lavelle[7], PhD; Sara Siddi[8,9,10], PhD; Femke Lamers[11], PhD; Brenda WJH Penninx[11], PhD; Josep Maria Haro[8,9,10], PhD; Peter Annas[12], PhD; Nicholas Cummins[1], PhD; Srinivasan Vairavan[13], PhD; Nikolay V Manyakov[2], PhD; Vaibhav A Narayan[13], PhD; Richard JB Dobson[1,3,4,5,6], PhD; Matthew Hotopf[4,7], PhD; RADAR-CNS[14]

[1]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[2]Data Science Analytics & Insights, Janssen Research & Development, Beerse, Belgium

[3]Institute of Health Informatics, University College London, London, United Kingdom

[4]NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, London, United Kingdom

[5]Health Data Research UK London, University College London, London, United Kingdom

[6]NIHR Biomedical Research Centre at University College London Hospitals NHS Foundation Trust, London, United Kingdom

[7]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[8]Teaching Research and Innovation Unit, Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain

[9]Centro de Investigación Biomédica, Red de Salud Mental, Madrid, Spain

[10]Faculty of Medicine and Health Sciences, Universitat de Barcelona, Barcelona, Spain

[11]Department of Psychiatry, Amsterdam Public Health Research Institute and Amsterdam Neuroscience, Amsterdam University Medical Centre, Vrije Universiteit and GGZ InGeest, Amsterdam, Netherlands

[12]H. Lundbeck A/S, Copenhagen, Denmark

[13]Janssen Research & Development, LLC, Titusville, NJ, United States

[14]see Acknowledgments

[*]these authors contributed equally

Corresponding Author:
Petroula Laiou, PhD
Department of Biostatistics and Health Informatics
Institute of Psychiatry, Psychology and Neuroscience
King's College London
Memory Lane
London, SE5 8AF
United Kingdom
Phone: 44 20 7848 0002
Email: petroula.laiou@kcl.ac.uk

Abstract

Background: Most smartphones and wearables are currently equipped with location sensing (using GPS and mobile network information), which enables continuous location tracking of their users. Several studies have reported that various mobility metrics, as well as home stay, that is, the amount of time an individual spends at home in a day, are associated with symptom severity in people with major depressive disorder (MDD). Owing to the use of small and homogeneous cohorts of participants, it is uncertain whether the findings reported in those studies generalize to a broader population of individuals with MDD symptoms.

XSL•FO
RenderX

Objective: The objective of this study is to examine the relationship between the overall severity of depressive symptoms, as assessed by the 8-item Patient Health Questionnaire, and median daily home stay over the 2 weeks preceding the completion of a questionnaire in individuals with MDD.

Methods: We used questionnaire and geolocation data of 164 participants with MDD collected in the observational Remote Assessment of Disease and Relapse–Major Depressive Disorder study. The participants were recruited from three study sites: King's College London in the United Kingdom (109/164, 66.5%); Vrije Universiteit Medisch Centrum in Amsterdam, the Netherlands (17/164, 10.4%); and Centro de Investigación Biomédica en Red in Barcelona, Spain (38/164, 23.2%). We used a linear regression model and a resampling technique (n=100 draws) to investigate the relationship between home stay and the overall severity of MDD symptoms. Participant age at enrollment, gender, occupational status, and geolocation data quality metrics were included in the model as additional explanatory variables. The 95% 2-sided CIs were used to evaluate the significance of model variables.

Results: Participant age and severity of MDD symptoms were found to be significantly related to home stay, with older (95% CI 0.161-0.325) and more severely affected individuals (95% CI 0.015-0.184) spending more time at home. The association between home stay and symptoms severity appeared to be stronger on weekdays (95% CI 0.023-0.178, median 0.098; home stay: 25th-75th percentiles 17.8-22.8, median 20.9 hours a day) than on weekends (95% CI −0.079 to 0.149, median 0.052; home stay: 25th-75th percentiles 19.7-23.5, median 22.3 hours a day). Furthermore, we found a significant modulation of home stay by occupational status, with employment reducing home stay (employed participants: 25th-75th percentiles 16.1-22.1, median 19.7 hours a day; unemployed participants: 25th-75th percentiles 20.4-23.5, median 22.6 hours a day).

Conclusions: Our findings suggest that home stay is associated with symptom severity in MDD and demonstrate the importance of accounting for confounding factors in future studies. In addition, they illustrate that passive sensing of individuals with depression is feasible and could provide clinically relevant information to monitor the course of illness in patients with MDD.

(JMIR Mhealth Uhealth 2022;10(1):e28095) doi:10.2196/28095

KEYWORDS

major depressive disorder; PHQ-8; smartphone; GPS; home stay; mobile phone

Introduction

The World Health Organization ranks depression as the single largest contributor to global disability [1]. People with major depressive disorder (MDD) often experience physical comorbidity [2], loss of occupational function [3], and low quality of life [4]. Furthermore, MDD is strongly associated with suicidal deaths and premature mortality [5]. The process for MDD diagnosis and evaluation of symptom severity is highly dependent on the subjective information that an individual under screening provides to a clinician, and it might be affected by recall bias.

Recent advances in digital technologies, including smartphones and wearable devices, enable the collection of a variety of data streams that can be used to objectively characterize an individual's daily activity and physical condition [6]. These data can be collected continuously, remotely, and unobtrusively without affecting an individual's daily routine and behavior. Importantly, analysis of such data could result in the development of new objective, quantifiable, cost-effective, and viable digital biomarkers of an individual's behavioral, cognitive, and emotional states [7-9]. Once developed and thoroughly tested, digital biomarkers hold great promise for improving the diagnosis and prognosis of a variety of mental health disorders, for facilitating continuous monitoring of individual well-being, and for supporting initiatives in precision medicine by helping to establish digital patient phenotypes in different disease areas [10].

Several recent studies have demonstrated the association between MDD symptoms and mobility patterns derived from mobile devices. For example, individuals with greater severity of MDD symptoms were reported to make fewer transitions between locations of interest (ie, those frequently visited in the past) and spend more time at home [11-15]. Home stay, an indicator of social disengagement [12], has also been reported to be significantly related to the severity of MDD symptoms [12-14].

Most studies used small and homogeneous cohorts of participants (eg, university students) and were conducted over a short period (eg, several weeks). In this study, we examined the association between the overall severity of MDD symptoms and a measure of daily mobility patterns using data from a larger and more diverse group of participants collected in the Remote Assessment of Disease and Relapse–Major Depressive Disorder (RADAR-MDD) study [16]. The RADAR-MDD study is an observational, longitudinal, prospective study that is currently being conducted at multiple clinical sites spread across several European countries and is part of a wider research program (Remote Assessment of Disease and Relapse–Central Nervous System [17]) to explore the potential of wearable devices to help prevent and treat depression, multiple sclerosis, and epilepsy. We used the 8-item Patient Health Questionnaire (PHQ-8; [18]) total score to measure the severity of MDD symptoms and home stay to describe an individual's daily mobility pattern. Home stay was selected as an interpretable measure of mobility with previous evidence suggesting that it is related to the severity of MDD symptoms [12-14]. In addition, we examined whether the strength of the relationship changes from weekdays to weekends (ie, modulated by changes in daily routine) and can be affected by an individual's demographics and quality of the acquired GPS recordings. We hypothesized

XSL•FO
RenderX

that higher levels of MDD, as quantified by the PHQ-8, would correspond to a prolonged home stay. In addition, we anticipated that the relationship between the severity of MDD symptoms and home stay would be modulated by changes in daily routine from weekdays to weekends [14]. If the hypotheses were proved to be true, this would provide additional evidence on the use of geolocation data in digital phenotyping [10].

Methods

Study Population

Participants were recruited for the RADAR-MDD study from November 2017 to November 2019. The recruited participants were aged ≥18 years and had experienced at least two episodes of MDD in their lifetime, with the most recent episode occurring within the last 2 years. The exclusion criteria included lifetime history of bipolar disorder; schizophrenia; MDD with psychotic features, schizoaffective disorders; history of moderate to severe drug or alcohol dependence within 6 months before enrollment; history of a major medical disease that could affect the participant's ability to be involved in normal daily activities for >2 weeks; dementia; and pregnancy. No limitations were applied regarding any treatment that the participants were receiving over the course of the study. Written consent was obtained before the enrollment session, followed by collection of sociodemographic, social environment, and medical history and technology use questionnaires and the Lifetime Depression Assessment Self-Report [19]. Participants with MDD were recruited from three clinical sites: King's College London (KCL) in the United Kingdom; Vrije Universiteit Medisch Centrum (VUMC) in Amsterdam, the Netherlands; and Centro de Investigación Biomédica en Red (CIBER) in Barcelona, Spain (Multimedia Appendix 1, Table S1 and Table 1). More details on the study protocol can be found in the study by Matcham et al [16].

Table 1. Data set characteristics (N=164).

Characteristic	Clinical site			
	KCL[a]	CIBER[b]	VUMC[c]	All sites
Participants with both PHQ-8[d] and GPS data collected, n (%)	232 (57.9)	116 (28.9)	53 (13.2)	401 (100)
Participants with biweekly segments fulfilling the selection criteria, n (%)	109 (66.4)	38 (23.2)	17 (10.4)	164 (100)
Female, n (%)	83 (76.1)	26 (68.4)	14 (82.4)	123 (75)
Age (years), median (range; SD)	46 (18-73; 15.0)	54 (27-71; 9.8)	33 (19-69; 14.9)	48 (18-73; 14.7)
Biweekly segments analyzed, n (%)	483 (62.8)	222 (28.9)	64 (8.3)	769 (100)
For employed participants	277 (72.5)	64 (16.8)	41 (10.7)	382 (100)
For unemployed participants	204 (53)	158 (41)	23 (6)	385 (100)

[a]KCL: King's College London.

[b]CIBER: Centro de Investigación Biomédica en Red.

[c]VUMC: Vrije Universiteit Medisch Centrum.

[d]PHQ-8: 8-item Patient Health Questionnaire.

The number of participants and biweekly segments collected at each site was normalized by the corresponding total obtained by pooling data across the 3 sites (column *All sites* in Table 1), with the resulting percentages indicated in parentheses. In all, 2 biweekly segments from KCL had no data on occupational status.

Data Collection

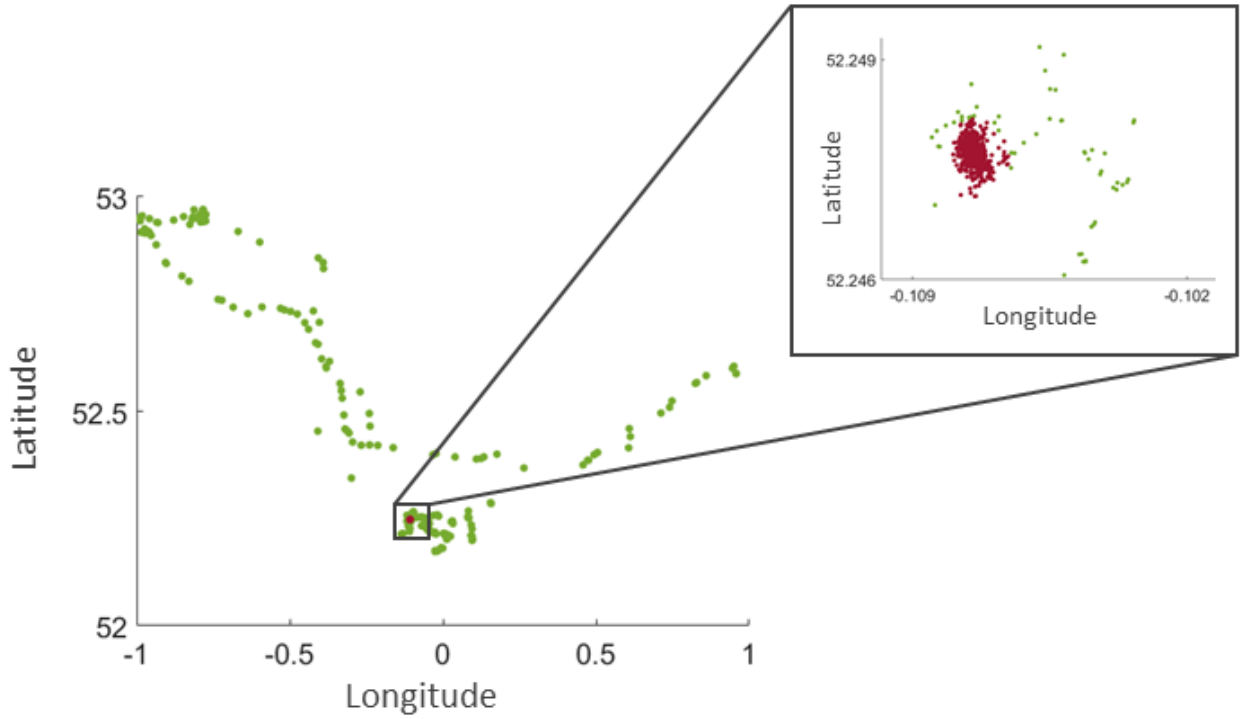
We used the RADAR-based platform for data collection and storage [20,21]. Participants with MDD were required to install several apps on their smartphones. Participants without a smartphone or with a non-Android device were provided with an Android smartphone and were required to use it throughout the study [16]. Remote monitoring technology (RMT) apps were used to collect data on participant severity of experienced MDD symptoms, self-esteem, cognitive functioning, voice audio sampling, and brief in-the-moment assessments of daily life experiences. Specifically, every 2 weeks, the participants were requested to fill in the PHQ-8 in the RADAR-base active RMT app. The request notifications were sent out at a calendared time and remained active only on that day initially. The completion window was increased to 3 days to improve the completion rates in April 2019.

The passive RMT apps ran in the background and required minimal input from the participants. The apps collected data on participants' physical (eg, transitions in space) and socially relevant activity (eg, number and duration of phone calls) as well as on some ambient factors (eg, ambient noise and light). GPS location data were obfuscated by adding a fixed random number to the latitude and longitude of all GPS data points generated by a single participant (Figure 1). The accuracy of each acquired GPS data point, as provided by either a mobile network operator or GPS satellites, corresponded to 1 SD (ie, radius) of a bivariate normal distribution with equal variances along the 2 spatial dimensions centered at that point. GPS data points with an accuracy >20 meters were discarded from the analyses. This accuracy level allowed the inclusion and analysis of most of the generated biweekly segments while ensuring high accuracy of the GPS recordings (Multimedia Appendix 1,

XSL•FO
RenderX

Table S2). The sampling period of the GPS signal was set to either 5 minutes [13,14,22] or 10 minutes [23-25] throughout the study. However, the effective sampling period varied over time towing to occasional signal loss and battery drain (ie, signal undersampling). Other factors affecting the sampling period included occasional concurrent and asynchronous acquisition of geolocation data from both a mobile network operator and GPS satellites (ie, signal oversampling).

Figure 1. Exemplar geolocation data which correspond to a biweekly segment of a study participant. The red dots denote individual's home location, whereas longitude and latitude along the axes are expressed in decimal degrees.



A single completed PHQ-8 combined with the GPS data acquired over the 2 preceding weeks and obtained from the same participant is herein referred to as a *biweekly segment*. To ensure a high quality of the analyzed geolocation data, only biweekly segments that met the following criteria were analyzed: 14 days of GPS recordings available, a daily median sampling period of the GPS signal ≤11 minutes, and the daily number of acquired GPS data points ≥48. The cutoff values were selected to maximize the volume of data available for analysis while preserving the high quality of these data (Multimedia Appendix 1, Tables S2-S4). Participants who declared their occupational status as a volunteer, student, caregiver, full- or part-time employee, or self-employed person were considered *employed*, whereas all other participants were considered *unemployed*. To avoid interference from the effect of the COVID-19 pandemic, only data collected before January 1, 2020, were analyzed.

For each day in a biweekly segment, we computed the number of GPS data points collected for each of the 24 hours separately and over the entire day. Ideally, a GPS signal sampled uniformly over a period of 5 minutes would give 12 GPS data points per hour and a total of 288 GPS data points per day. We specified completeness of the daily data as a ratio between the actual number of GPS data points collected over a day and the expected number as determined by a sampling period (ie, 288 and 144 GPS data points for the sampling periods of 5 minutes and 10 minutes, respectively). The extreme completeness values of 0.0 and 1.0 correspond to an empty and a complete day of GPS recordings, respectively, with the values in-between corresponding to partial or interrupted GPS recordings throughout a day (Multimedia Appendix 1, Table S5). Multimedia Appendix 1, Figure S1 shows the median data completeness for each hour in a day across the analyzed biweekly segments. Similarly, we divided the actual number of GPS data points collected per hour by the expected number and computed the SD of these 24 normalized values to characterize the sampling constancy of the daily data. Any positive value for sampling constancy indicates fluctuations in the volume of GPS data acquired throughout the day, with greater values indicating greater fluctuations (Multimedia Appendix 1, Table S6). Median completeness and median sampling constancy of the daily data, as computed across 14 days of a biweekly segment, were used to characterize the quality of GPS recordings acquired for that segment.

Home Stay

Home location was identified in a stepwise manner. Initially, the home location was approximated by the median longitude and latitude of all GPS data points in a biweekly segment acquired between 12 AM and 6 AM [12-14]. To account for accidental travel and outdoor stay, all GPS data points acquired during these hours and separated by >60 meters (=3 × the accuracy level of ≤20 meters; see the *Data Collection* section) from the initial home location were discarded. The home location was finally determined as the median longitude and latitude of all remaining GPS data points (Figure 1). The distance between any 2 GPS data points *i* and *j* was computed using the Haversine WGS84 formula as follows [22]:

XSL•FO
RenderX

$$\text{distance} = \sin^2(\Delta\varphi/2) + \cos\varphi i \cos\varphi j \sin^2(\Delta\lambda/2)$$

where φ and λ correspond to the latitude and longitude, respectively.

The home stay for a given day was specified as the ratio between the number of GPS data points separated by ≤60 meters from the home location and the total number of GPS data points acquired on that day. Home stay values of 0 (or 0%) and 1 (or 100%) correspond to an entire day spent outside versus at home, respectively. Median home stay, as computed across 14 days of a biweekly segment, was used to characterize the home stay of a study participant for that segment (Multimedia Appendix 1, Table S7).

Statistical Analysis

A linear regression model was selected to test the relationship between home stay and overall severity of MDD symptoms. Specifically, home stay was used as a dependent variable, with PHQ-8 total score being used as an independent variable. Participant age at enrollment, gender (men vs women), occupational status (employed vs unemployed), median completeness, and sampling constancy of the daily data in a biweekly segment were included in the model as additional explanatory variables:

home stay≈PHQ-8 score+age+gender+occupational status+data completeness+sampling constancy

We chose home stay as a dependent variable to test its relationship not only to the severity of MDD symptoms but also to participants' demographics and quality characteristics of the collected geolocation data in a single model in a uniform manner. For each study participant, we randomly selected one of the biweekly segments generated by that participant. The model was fitted using data from biweekly segments pooled across the participants. To obtain a CI for each of the 6 regression coefficients, the procedure of random selection of a biweekly segment per participant followed by pooling data across the participants and fitting the model was repeated 100 times. A model variable was deemed to be significantly related to home stay if a 95% 2-sided CI obtained for the regression coefficient of that parameter did not include 0. The model was fitted using data from all 3 sites combined (Table 2) and each clinical site separately (Multimedia Appendix 1, Table S8).

Table 2. CIs and medians for the 6 regression coefficients of the linear regression model.

Analyzed time frame	Value, median (95% CI)					
	Age	Gender	PHQ-8[a] total score	Occupational status	Median completeness of the daily data	Median sampling constancy of the daily data
Over the entire week	*0.241 (0.161 to 0.325)* [b]	–0.121 (–0.272 to 0.024)	*0.100 (0.015 to 0.184)*	*–0.448 (–0.631 to –0.279)*	–0.044 (–0.108 to 0.022)	–0.064 (–0.130 to 0.005)
Weekdays only	*0.254 (0.188 to 0.329)*	–0.061 (–0.220 to 0.058)	*0.098 (0.023 to 0.178)*	*–0.495 (–0.664 to –0.354)*	–0.024 (–0.097 to 0.034)	–0.041 (–0.110 to 0.041)
Weekends only	*0.148 (0.029 to 0.240)*	–0.036 (–0.265 to 0.174)	0.052 (–0.079 to 0.149)	*–0.323 (–0.535 to –0.127)*	0.023 (–0.074 to 0.101)	–0.075 (–0.144 to 0.006)

[a]PHQ-8: 8-item Patient Health Questionnaire.

[b]CIs that do not include 0 are italicized. The regression coefficients obtained with standardized data for each clinical site separately are reported in Multimedia Appendix 1, Table S8.

The model was fitted using standardized data pooled across the 3 sites for each analyzed time frame separately. The positive sign of the regression coefficients that correspond to the categorical variables (ie, gender and occupational status) indicates greater home stay for men and employed as compared with women and unemployed participants, respectively. All reported CIs are 95% 2-sided intervals.

To test whether the relationship between home stay and the independent variables differed between weekdays and weekends, a similar approach was followed. Specifically, home stay, median completeness, and sampling constancy of the daily data in a biweekly segment were estimated separately for weekdays and weekends. As a single biweekly segment included 10 weekdays and only 4 weekend days, we used 4 days to generate those estimates to equalize variance in the estimates of weekdays and weekends. For each analyzed biweekly segment, we randomly drew 4 weekdays 100 times. The medians of the estimates computed for each of these 100 draws were used to characterize the weekdays of that segment in the model.

To account for nonnormality of both dependent and independent variables as well as for differences in their variance, each variable (except for gender and occupational status) was standardized by applying the Yeo–Johnson transformation followed by the zero-mean, unit-variance normalization. All models and findings reported throughout the manuscript were obtained by using these *standardized* data. However, qualitatively similar results were obtained when using the original, nonstandardized data (Multimedia Appendix 1, Table S9). All statistical analyses were performed using the Matrix Laboratory R2019b.

Results

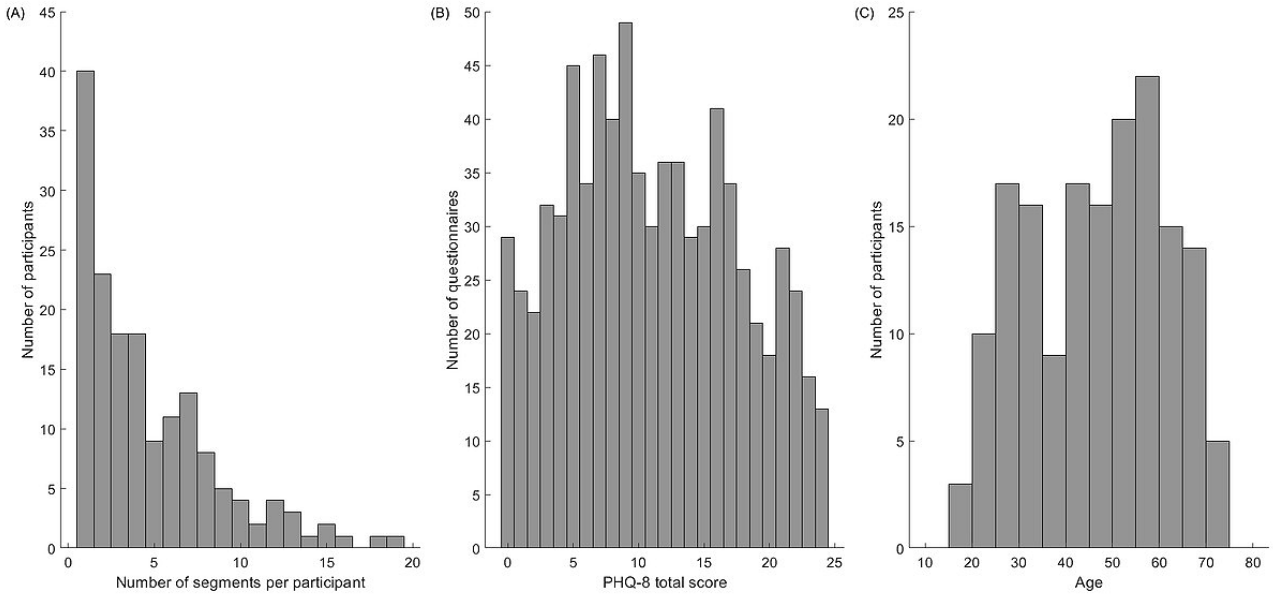
Data Set Characteristics

As of January 1, 2020, the total number of participants enrolled in the RADAR-MDD study across the 3 clinical sites was 432 (Multimedia Appendix 1, Table S1). Of those 432, a total of 401 (92.8%) participants had usable PHQ-8 and geolocation data (Table 1), resulting in a total of 4273 biweekly segments

XSL•FO
RenderX

generated across the sites (Multimedia Appendix 1, Table S2). After discarding GPS data points with low accuracy (>20 meters) and selecting only biweekly segments with 14 days of GPS recordings available, the number of biweekly segments reduced to 43.9% (1876/4273; Multimedia Appendix 1, Table S2). Imposing additional requirements on the daily median sampling period (≤11 minutes; Multimedia Appendix 1, Table S3) and the daily minimum volume (≥48 data points; Multimedia Appendix 1, Table S4) of the GPS data in a single biweekly segment further reduced the number of biweekly segments available for analysis to 17.99% (769/4273; Table 1). The latter corresponds to data from 38% (164/432) of study participants. Table 1 lists the demographic characteristics of the participants, whereas Figure 2C shows the distribution of their ages. Most of the study participants enrolled at each clinical site were women (range 26/38, 68%-14/17, 82%; Table 1; [26]).

Figure 2. Distributions of data set characteristics. (A) Number of biweekly segments available for analysis per study participant. (B) 8-item Patient Health Questionnaire total score. (C) Participant age. Data were pooled across the 3 clinical sites. PHQ-8: 8-item Patient Health Questionnaire.



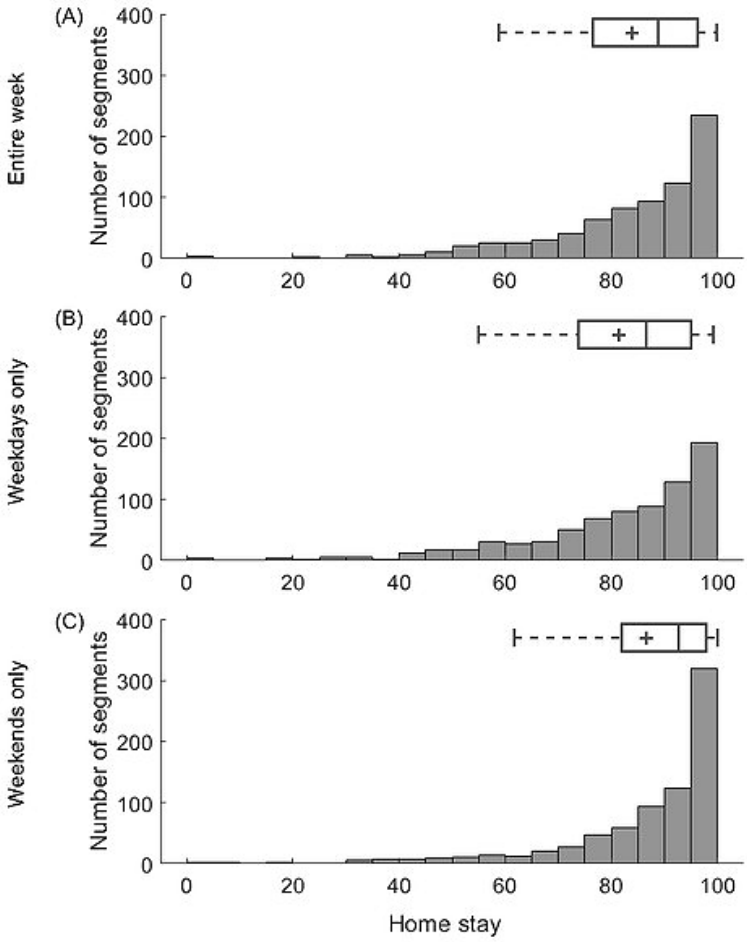
The number of biweekly segments available for analysis varied considerably across the sites, with VUMC (64/769, 8.3% segments; Table 1) and KCL (483/769, 62.8%) providing the least and most data, respectively. The number of biweekly segments produced by a single participant varied between 1 and 19, with a median equaling 4 (25th-75th percentiles 2-7; Figure 2A). As shown in Figure 2B, the collected data represented all 5 severity categories of MDD, as specified in the PHQ-8 questionnaire [18], including *none-minimal* (PHQ-8 total score from 0 to 4; 138/769, 18% segments), *mild* (PHQ-8 total score from 5 to 9; 214/769, 27.8% of the total), *moderate* (PHQ-8 total score from 10 to 14; 166/769, 21.6% of the total), *moderately severe* (PHQ-8 total score from 15 to 19; 152/769, 19.8% of the total), and *severe* (PHQ-8 total score from 20 to 24; 99/769, 12.8% of the total). The data set characteristics for each individual site and occupational status are shown in Multimedia Appendix 1, Figures S2 and S3.

Estimates of Home Stay

Over the course of the study, the participants spent most of their time at home. When no distinction between weekdays and weekends was made, median home stay across the sites was 89% (21.4 hours a day; 25th-75th percentiles 76%-96% or 18.2-23.0 hours a day; Figure 3A). As expected, the home stay was lower during weekdays than during the weekends (Figures 3B and 3C). Specifically, the median home stay across the sites was 87% (20.9 hours a day; 25th-75th percentiles 74%-95% or 17.8-22.8 hours a day) and 93% (22.3 hours a day; 25th-75th percentiles 82%-98% or 19.7-23.5 hours a day) when analyzing weekday and weekend data, respectively. These observations were consistent across each clinical site (Multimedia Appendix 1, Figure S4 and Table S7).

XSL•FO
RenderX

Figure 3. Home stay computed (A) over the entire week, (B) for weekdays, and (C) weekends only. A grey horizontal bar and a cross in each boxplot indicate median and mean of the presented data. Data were pooled across the 3 clinical sites. Home stay corresponds to the percentage of time spent at home a day.



Similarly, home stay was affected by occupational status. The employed participants spent less time at home compared with their unemployed peers. Median home stay across the sites was 82% (19.7 hours a day; 25th-75th percentiles 67%-92% or 16.1-22.1 hours a day) and 94% (22.6 hours a day; 25th-75th percentiles 85%-98% or 20.4-23.5 hours a day) for the employed and unemployed participants, respectively, with the difference being more prominent during weekdays (79% vs 93% or 19.0 vs 22.3 hours a day) than during the weekends (88% vs 96% or 21.1 vs 23.0 hours a day). The same pattern of observations was seen across each clinical site (Multimedia Appendix 1, Figure S5 and Table S7).

Associations With Home Stay

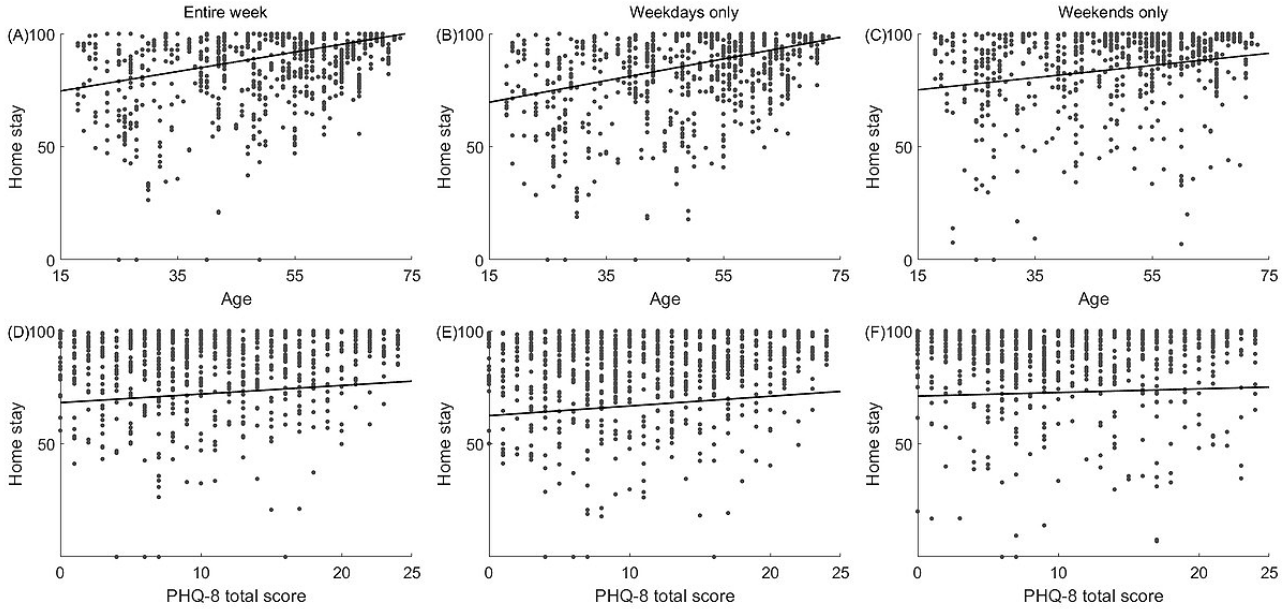
When data were pooled across the sites and no distinction between weekdays and weekends was made, the linear regression model revealed a significant relationship between home stay and overall severity of the MDD symptoms as captured by the PHQ-8 total score (median 0.100, 2-sided 95% CI 0.015-0.184; Figure 4D; Table 2). The latter suggested that greater overall severity of MDD symptoms was associated with prolonged home stay. The same relationship was observed when analyzing weekday data only (median 0.098, 95% CI 0.023-0.178; Figure 4E) but not on weekends (median 0.052, 95% CI −0.079 to 0.149; Figure 4F).

In addition, the model revealed a significant relationship between home stay and age. Specifically, the participants spent more time at home with age (median 0.241, 95% CI 0.161-0.325; Figure 4A; Table 2). A similar strength of the relationship was observed for weekdays (median 0.254, 95% CI 0.188-0.329; Figure 4B) and weekends (median 0.148, 95% CI 0.029-0.240; Figure 4C). Furthermore, occupational status was also found to significantly modulate home stay, with the employed participants spending less time at home compared with their unemployed peers (median −0.448, 95% CI −0.631 to −0.279; Table 2). Similar to age, there was no significant difference in the effect of occupational status on home stay among the analyzed time frames (weekdays only: median −0.495, 95% CI −0.664 to −0.354; weekends only: median −0.323, 95% CI −0.535 to −0.127).

Neither gender nor median completeness and sampling constancy of the daily data in a biweekly segment had a significant impact on home stay and this held for all the analyzed time frames (Table 2). The results of modeling for each clinical site obtained with standardized and original data are shown in Multimedia Appendix 1, Tables S8 and S9, respectively.

XSL•FO
RenderX

Figure 4. Relationship (A, B, and C) between home stay and participant age and (D, E, and F) between home stay and the 8-item Patient Health Questionnaire total score as assessed using data of (A and D) the entire week, (B and E) weekdays, and (C and F) weekends only. Each dot indicates a single biweekly segment. Data of all biweekly segments pooled across the 3 clinical sites are presented. A black line in each panel corresponds to the linear fit of the presented data. Home stay corresponds to the percentage of time spent at home a day. PHQ-8: 8-item Patient Health Questionnaire.



Discussion

Principal Findings

Multiple studies have demonstrated associations between patterns of daily movements of an individual in an area of the primary residence and an individual's mood [9]. Here, we tested the association between home stay and overall severity of MDD symptoms, as reflected in the PHQ-8 total score, by using data collected in the RADAR-MDD study. The participants were invited to complete the PHQ-8 on their mobile phones every 2 weeks, whereas the same phones were used to track their geographic location continuously throughout the study. We related the PHQ-8 total score, as provided by an individual, to their median daily home stay over the 2 weeks preceding completion of the PHQ-8. In addition, we investigated how the relationship between home stay and MDD symptom severity was affected by participant age, gender, occupational status as well as by completeness and sampling constancy of the collected geolocation data. Moreover, we tested whether the strength of the relationship differed between weekdays and weekends.

The participants in the RADAR-MDD study were recruited from a nonhomogeneous population (ie, clinical and community samples with a wide age range) across 3 clinical sites in different European countries. When we pooled the data from all sites and used the entire biweekly segment before PHQ-8 completion, we found that home stay was positively associated with the PHQ-8 total score and age (Table 2). Specifically, the participants tended to spend more time at home with a greater severity of MDD symptoms and age. Furthermore, we found that occupational status was significantly related to home stay, with unemployed participants spending more time at home than their employed peers. Similar findings were observed when analyzing geolocation data collected over weekdays or weekends only, except for the association between home stay and the PHQ-8 total score (Table 2). The latter failed to reach statistical significance when tested with geolocation data of weekends only. This can be attributed to the ceiling effect [27], as the estimates of home stay obtained with geolocation data of the weekends were high for almost all participants at each individual site (Figure 3C; Multimedia Appendix 1, Figures S4 and S5 and Table S7). Although similar findings were observed for the KCL and CIBER sites, the association between home stay and the PHQ-8 total score did not reach statistical significance at the latter site (Multimedia Appendix 1, Table S8). This discrepancy could have been driven by participant recruitment primarily in a clinical setting and right skew of the PHQ-8 total scores indicating great severity of the MDD symptoms in the participants recruited at the CIBER site (Multimedia Appendix 1, Figure S2). As the VUMC site recruited only 17 participants (Table 1) that, on average, exhibited mild to moderate symptoms of MDD (Multimedia Appendix 1, Figure S2), all the findings obtained with the data of that site only should be interpreted with caution.

Comparison With Previous Work

A variety of features can be extracted from geolocation data generated by smartphones and wearable devices and used to characterize the mobility patterns of an individual. These include home stay [15], the number of visited places [14], location entropy (ie, a metric that quantifies uniformity of the distribution of times spent by an individual at different locations) [13,14], the maximal distance from home, and the total distance traveled [22]. Remarkably, several studies that investigated the relationship between mental health disorders and mobility patterns focused on home stay features [12,22,28]. In this study, we also used home stay to quantify the mobility patterns of the

XSL•FO
RenderX

study participants, as home stay is considered an important indicator of social disengagement by clinicians [12]. Moreover, it has been demonstrated that home stay has a strong negative association with location entropy [9]. No features that quantify the distance traveled between visited locations (eg, the total distance traveled or the maximal distance from home) were used in our analysis, as the notion of distance was confounded by the fact that the participants lived in both urban and rural places and in different countries.

Several previous studies have documented a positive relationship between home stay and the severity of MDD symptoms [9,13-15]. To the best of our knowledge, no study, however, has collected data from either multiple sites or a nonhomogeneous population with a confirmed clinical diagnosis of MDD. Neither did those studies thoroughly address the factors of participants' age, occupational status, and data quality on the reported results. Furthermore, several previous studies analyzed data that were homogeneous in terms of the participant's age (ie, student population) [12,24]. In contrast, the age of the RADAR-MDD participants ranged from 18 to 73 years (Figure 2C). Our findings demonstrate that the strength of the relationship between home stay and the severity of MDD symptoms can be modulated by age. This relationship is expected to be stronger for younger individuals and weaker for older individuals, as the latter tend to stay at home more. In addition, all RADAR-MDD participants had a clinical diagnosis of MDD, and many of them had severe symptoms of MDD, as indicated by the high PHQ-8 total scores (Figure 2B). In contrast, participants from previous studies did not undergo clinical interviews and had overall low depression scores [11-14].

It has been pointed out [9,22] that there exists no standard approach to the preprocessing of geolocation data generated by smartphones. Nonetheless, such important preprocessing steps, such as selection of an acceptable accuracy level and rate of missing data for the geolocation signal could have a significant impact on the reported results. In this study, we did not use all available geolocation data collected from the RADAR-MDD participants but instead applied stringent selection criteria (see *Data Collection*) to ensure high quality of the analyzed data and minimize the odds of reporting spurious results. In addition, we provide full and detailed information on the characteristics of the collected and analyzed geolocation data (Multimedia Appendix 1, Tables S1-S6).

Limitations

Although this was a multicenter study and the estimates of home stay were similar across the 3 sites, most participants were recruited at KCL (Multimedia Appendix 1, Figures S4 and S5 and Table S7). Several participants in the RADAR-MDD study followed antidepressant treatment, and some of them reported comorbidity with physical illness (eg, fibromyalgia). In addition, several participants were off sick or reported ill health. Antidepressants may cause a wide range of side effects, including headaches, fatigue, weight gain, drowsiness, and dizziness [29,30]. Individuals who experience any or all of these side effects or have comorbidities with physical illness are likely to spend more time at home than outdoors. This could have inflated the reported estimates of home stay (Figure 3) and thus distorted the strength of the relationship between home stay and overall severity of MDD symptoms.

Apart from medical and mental conditions, social factors may have also influenced how much time participants spent at home. These include the number of people living under the same roof and engaging in outdoor or community activities. The participants who were expected to assist their elderly family members in daily routines or take care of their children likely spent more time at home than their peers without such responsibilities. In contrast, engagement in outdoor or community activities, such as playing bingo or going to church, likely resulted in reduced home stay. Furthermore, it is commonly assumed that employment implies the physical presence of an employee in a designated workplace outside of home. However, we cannot rule out that some employed participants worked from home. Home teleworking likely increased home stay for those participants. As employment was significantly associated with reduced home stay in our data set, most employed participants in the study still worked outside their home. The effect of medication and physical comorbidity, social factors, and home teleworking on the relationship between daily mobility patterns and severity of MDD symptoms was beyond the scope of this study, although further research is warranted.

The stringent selection criteria imposed on completeness and sampling constancy of the collected geolocation data considerably reduced the number of biweekly segments available for analysis. Several factors could have affected the quality of the collected geolocation data. Poor mobile network coverage or weak GPS signals, for example, was expected to result in a higher missing rate of geolocation recordings. This was likely the case for participants living or traveling in distant or rural areas. Smartphone battery capacity could have constrained the total duration of the geolocation recordings. Owing to a limited battery capacity, frequent user interaction with a smartphone could have accelerated the battery drain and further limited the total duration of geolocation recordings. In addition, a high number of apps running in the background could have also contributed to a more rapid battery drain. The RADAR-MDD study was designed to concurrently collect a variety of data streams (eg, from a GPS sensor, a gyroscope, an accelerometer, a microphone, and an ambient light sensor embedded in a smartphone) to characterize the individual's behavior at full capacity [20,21]. This resulted in greater energy consumption and thus faster battery drain than in regular smartphones with no installed RADAR-MDD apps. Disabling the collection of one or multiple data streams in the study could have considerably prolonged the time smartphones operated on a single battery charge. Identification and comprehensive characterization of a single data stream or multiple data streams that convey most information on the individual's mental well-being is still a topic of active scientific research. Alternatively, event-driven collection of all or some data streams (eg, initiated by the individual's accidental movements or continuous motion) could have been less energy demanding than continuous sampling of those data streams as was implemented in the RADAR-MDD study. Finally, it is uncertain

XSL•FO
RenderX

and requires additional examination whether the same results could have been obtained with more liberal selection criteria (eg, at least 10 instead of 14 days of recordings in a biweekly segment). If so, this would have increased the number of biweekly segments available for analysis and provided stronger evidence in support of the feasibility of geolocation data collection with smartphones.

Conclusions

We demonstrated that longer home stay can reflect greater symptom severity in individuals diagnosed with MDD. Although the relationship between home stay and MDD severity is modest, it can nonetheless improve remote monitoring of the individual's mental well-being, especially when combined with other informative correlates of MDD severity. However, it remains unclear whether the findings represent behavioral manifestations of MDD or are associated with changes in depressive symptoms. Additional analyses are required to test whether changes in home stay over time can be predictive of relapses in MDD. We also demonstrated that the relationship between home stay and MDD severity can be modulated by age, occupational status, and changes in daily routine. This finding is of great importance for a proper interpretation of similar studies conducted in the past and for better planning of future studies. Furthermore, our findings illustrate that passive remote monitoring of mobility patterns in individuals with MDD is feasible. This demonstrates the utility of smartphones and wearable devices with a GPS sensor in the collection of clinically relevant information that can be used to monitor the course of the disorder in a remote, unobtrusive, and ubiquitous manner, thus reducing patient burden and improving treatment.

Acknowledgments

The Remote Assessment of Disease and Relapse–Central Nervous System project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement number 115902. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and European Federation of Pharmaceutical Industries and Associations [31]. This communication reflects the views of the Remote Assessment of Disease and Relapse–Central Nervous System consortium and neither Innovative Medicines Initiative nor the European Union and European Federation of Pharmaceutical Industries and Associations are liable for any use that may be made of the information contained herein. The funding body have not been involved in the design of the study, the collection or analysis of data, or the interpretation of data.

Participant recruitment in Amsterdam was partially accomplished through Hersenonderzoek.nl, a Dutch online registry that facilitates participant recruitment for neuroscience studies [32]. Hersenonderzoek.nl is funded by ZonMw-Memorabel (project number 73305095003), a project in the context of the Dutch Deltaplan Dementie, Gieskes-Strijbis Foundation, the Alzheimer's Society in the Netherlands and Brain Foundation Netherlands.

Participants in Spain were recruited through the following institutions: Parc Sanitari Sant Joan de Déu network of mental health services (Barcelona); Institut Català de la Salut primary care services (Barcelona); Institut Pere Mata-Mental Health Care (Tarrassa); Hospital Clínico San Carlos (Madrid).

This paper represents independent research part funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley National Health Service (NHS) Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

We thank all the members of the Remote Assessment of Disease and Relapse–Central Nervous System patient advisory board for their contribution to the device selection procedures, and their invaluable advice throughout the study protocol design.

This research was reviewed by a team with experience of mental health problems and their carers who have been specially trained to advise on research proposals and documentation through the Feasibility and Acceptability Support Team for Researchers: a free, confidential service in England provided by the National Institute for Health Research Maudsley Biomedical Research Centre via King's College London and South London and Maudsley NHS Foundation Trust.

We thank all Genetic Links to Anxiety and Depression Study volunteers for their participation, and gratefully acknowledge the NIHR BioResource, NIHR BioResource centres, NHS Trusts and staff for their contribution. We also acknowledge NIHR BRC, King's College London, South London and Maudsley NHS Trust and King's Health Partners. We thank the National Institute for Health Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme.

Authors' Contributions

PL and DAK extracted and integrated the questionnaire and geolocation data for the analysis, planned and performed the analysis, and drafted the manuscript. MH and VAN gained funding and co-led the Remote Assessment of Disease and Relapse–Central Nervous System program. MH is the principal investigator for the Remote Assessment of Disease and Relapse–Major Depressive Disorder study. RJBD, AAF, YR, ZR, PC, and CS have contributed to the development of the Remote Assessment of Disease and Relapse–based platform used for data collection and management across sites, data protection, security, and storage. PL, DAK, AAF, S Sun, YZ, FM, FL, S Siddi, BWJHP, JMH, PA, NC, SV, NVM, VAN, RJBD, and MH contributed to the design

XSL•FO
RenderX

of the study. FM, AI, GL, and S Siddi collected the data. All authors meet the International Committee of Medical Journal Editors criteria, and all those who fulfilled these criteria are listed as authors. All authors have been involved in reviewing the manuscript, had access to the study data, provided direction and comments on the manuscript, made the final decision about where to publish these data, and approved submission to this journal. All authors agree with the content and author list of this manuscript.

Conflicts of Interest

DAK and NVM are employees of the Janssen Pharmaceutica NV and may hold company equity. SV and VAN are employees of Janssen Research and Development, LLC, and may hold company equity. PA is an employee of H. Lundbeck A/S and may hold company equity. MH declares research grants and in-kind contributions from Janssen, Biogen, UCB, MSD, and H. Lundbeck A/S through the Remote Assessment of Disease and Relapse–Central Nervous System consortium.

Multimedia Appendix 1
Tables and figures with information for each individual population site.
[DOCX File , 509 KB - mhealth_v10i1e28095_app1.docx]

References

1. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/handle/10665/254610 [accessed 2021-12-30]
2. Cimpean D, Drake RE. Treating co-morbid chronic medical conditions and anxiety/depression. Epidemiol Psychiatr Sci 2011 Jun 08;20(2):141-150. [doi: 10.1017/s2045796011000345] [Medline: 21714361]
3. Lerner D, Adler DA, Chang H, Berndt ER, Irish JT, Lapitsky L, et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. J Occup Environ Med 2004 Jun;46(6 Suppl):46-55 [FREE Full text] [doi: 10.1097/01.jom.0000126684.82825.0a] [Medline: 15194895]
4. Lenox-Smith A, Macdonald MT, Reed C, Tylee A, Peveler R, Quail D, et al. Quality of llfe in depressed patients in UK primary care: the FINDER study. Neurol Ther 2013 Dec 13;2(1-2):25-42 [FREE Full text] [doi: 10.1007/s40120-013-0006-1] [Medline: 26000214]
5. Cuijpers P, Schoevers RA. Increased mortality in depressive disorders: a review. Curr Psychiatry Rep 2004 Dec;6(6):430-437. [doi: 10.1007/s11920-004-0007-y] [Medline: 15538991]
6. Callan JA, Wright J, Siegle GJ, Howland RH, Kepler BB. Use of computer and mobile technologies in the treatment of depression. Arch Psychiatr Nurs 2017 Dec;31(3):311-318. [doi: 10.1016/j.apnu.2016.10.002] [Medline: 28499574]
7. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
8. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017 May 08;13:23-47. [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
9. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e165 [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]
10. Liang Y, Zheng X, Zeng DD. A survey on big data-driven digital phenotyping of mental health. Inform Fusion 2019 Dec;52:290-307. [doi: 10.1016/j.inffus.2019.04.001]
11. Canzian L, Musolesi M. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp '15: The 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7 - 11, 2015; Osaka Japan p. 1293-1304. [doi: 10.1145/2750858.2805845]
12. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. J Med Internet Res 2017 Mar 03;19(3):e62 [FREE Full text] [doi: 10.2196/jmir.6820] [Medline: 28258049]
13. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
14. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
15. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
16. Matcham F, di San Pietro CB, Bulgari V, de Girolamo G, Dobson R, Eriksson H, RADAR-CNS Consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry 2019 Feb 18;19(1):72 [FREE Full text] [doi: 10.1186/s12888-019-2049-z] [Medline: 30777041]

XSL•FO
RenderX

17. RADAR-CNS: Remote Assessment of Disease and Relapse – Central Nervous System. RADAR-CNS. URL: http://radar-cns.org/ [accessed 2021-05-23]
18. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
19. Bot M, Middeldorp CM, de Geus EJ, Lau HM, Sinke M, van Nieuwenhuizen B, et al. Validity of LIDAS (LIfetime Depression Assessment Self-report): a self-report online assessment of lifetime major depressive disorder. Psychol Med 2016 Oct 05;47(2):279-289. [doi: 10.1017/s0033291716002312]
20. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Hyve, RADAR-CNS Consortium. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]
21. RADAR-base: Remote Assessment of Disease and Relapse - an mHealth Platform for Generalized Data Collection. RADAR-base. URL: http://radar-base.org/ [accessed 2021-05-23]
22. Depp CA, Bashem J, Moore RC, Holden JL, Mikhael T, Swendsen J, et al. GPS mobility as a digital biomarker of negative symptoms in schizophrenia: a case control study. NPJ Digit Med 2019 Nov 08;2(1):108 [FREE Full text] [doi: 10.1038/s41746-019-0182-1] [Medline: 31728415]
23. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
24. Farhan A, Lu J, Bi J, Russell A, Wang B, Bamis A. Multi-view bi-clustering to identify smartphone sensing features indicative of depression. In: In Proceedings of the IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). 2016 Presented at: IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE); June 27-29, 2016; Washington, DC, USA p. 264-273. [doi: 10.1109/chase.2016.27]
25. Farhan A, Yue C, Morillo R, Ware S, Lu J, Bi J, et al. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. In: Proceedings of the IEEE Wireless Health (WH) Conference. 2016 Presented at: IEEE Wireless Health (WH) Conference; Oct. 25-27, 2016; Bethesda, MD, USA p. 1-8. [doi: 10.1109/wh.2016.7764553]
26. Albert P. Why is depression more prevalent in women? J Psychiatry Neurosci 2015 Jul 1;40(4):219-221 [FREE Full text] [doi: 10.1503/jpn.150205] [Medline: 26107348]
27. Šimkovic M, Träuble B. Robustness of statistical methods when measure is affected by ceiling and/or floor effect. PLoS One 2019 Aug 19;14(8):e0220889 [FREE Full text] [doi: 10.1371/journal.pone.0220889] [Medline: 31425561]
28. Tung JY, Rose RV, Gammada E, Lam I, Roy EA, Black SE, et al. Measuring life space in older adults with mild-to-moderate Alzheimer's disease using mobile phone GPS. Gerontology 2014;60(2):154-162. [doi: 10.1159/000355669] [Medline: 24356464]
29. Ferguson JM. SSRI antidepressant medications: adverse effects and tolerability. Prim Care Compan J Clin Psychiatry 2001 Mar 01;3(1):22-27 [FREE Full text] [doi: 10.4088/pcc.v03n0105] [Medline: 15014625]
30. Wang S, Han C, Bahk W, Lee S, Patkar AA, Masand PS, et al. Addressing the side effects of contemporary antidepressant drugs: a comprehensive review. Chonnam Med J 2018 May;54(2):101-112 [FREE Full text] [doi: 10.4068/cmj.2018.54.2.101] [Medline: 29854675]
31. Europe's partnership for health. Innovative Health Initiative. URL: https://www.imi.europa.eu/ [accessed 2022-01-11]
32. Hersenonderzoek. URL: https://hersenonderzoek.nl/ [accessed 2022-01-11]

Abbreviations

CIBER: Centro de Investigación Biomédica en Red
KCL: King's College London
MDD: major depressive disorder
PHQ-8: 8-item Patient Health Questionnaire
RADAR-MDD: Remote Assessment of Disease and Relapse–Major Depressive Disorder
RMT: remote monitoring technology
VUMC: Vrije Universiteit Medisch Centrum

XSL•FO
RenderX

Edited by L Buis; submitted 21.02.21; peer-reviewed by A Baglione, A AL-Asadi, D Mack, W Chan; comments to author 27.04.21; revised version received 20.06.21; accepted 21.10.21; published 28.01.22.

Please cite as:
Laiou P, Kaliukhovich DA, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, Sun S, Zhang Y, Matcham F, Ivan A, Lavelle G, Siddi S, Lamers F, Penninx BWJH, Haro JM, Annas P, Cummins N, Vairavan S, Manyakov NV, Narayan VA, Dobson RJB, Hotopf M, RADAR-CNS
The Association Between Home Stay and Symptom Severity in Major Depressive Disorder: Preliminary Findings From a Multicenter Observational Study Using Geolocation Data From Smartphones
JMIR Mhealth Uhealth 2022;10(1):e28095
URL: https://mhealth.jmir.org/2022/1/e28095
doi:10.2196/28095
PMID:35089148

©Petroula Laiou, Dzmitry A Kaliukhovich, Amos A Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Shaoxiong Sun, Yuezhou Zhang, Faith Matcham, Alina Ivan, Grace Lavelle, Sara Siddi, Femke Lamers, Brenda WJH Penninx, Josep Maria Haro, Peter Annas, Nicholas Cummins, Srinivasan Vairavan, Nikolay V Manyakov, Vaibhav A Narayan, Richard JB Dobson, Matthew Hotopf, RADAR-CNS. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 28.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile App to Increase Fruit and Vegetable Acceptance Among Finnish and Polish Preschoolers: Randomized Trial

Henna Vepsäläinen[1], PhD; Essi Skaffari[1], MSc; Katarzyna Wojtkowska[2], PhD; Julia Barlińska[2], PhD; Satu Kinnunen[1], MSc; Riikka Makkonen[1,3], MSc; Maria Heikkilä[1], PhD; Mikko Lehtovirta[4], MD, PhD; Carola Ray[1,5], PhD; Eira Suhonen[6], PhD; Jaakko Nevalainen[7], PhD; Nina Sajaniemi[3,6], PhD; Maijaliisa Erkkola[1], PhD

[1]Department of Food and Nutrition, University of Helsinki, Helsinki, Finland
[2]Faculty of Psychology, University of Warsaw, Warsaw, Poland
[3]School of Applied Educational Science and Teacher Education, University of Eastern Finland, Joensuu, Finland
[4]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland
[5]Folkhälsan Research Center, Helsinki, Finland
[6]Department of Education, University of Helsinki, Helsinki, Finland
[7]Faculty of Social Sciences, Tampere University, Tampere, Finland

Corresponding Author:
Henna Vepsäläinen, PhD
Department of Food and Nutrition
University of Helsinki
PO Box 66
Helsinki, 00014
Finland
Phone: 358 443581467
Email: henna.vepsalainen@helsinki.fi

Abstract

Background: Early childhood education and care (ECEC) centers are ideal venues for food education. As smartphones and tablets are becoming increasingly popular in ECEC centers, technology can be used to deliver such pedagogical content. Evidence suggests that video games can affect fruit and vegetable (FV) consumption among 9- to 12-year-old children, but studies among preschoolers are scarce.

Objective: This paper describes the development of the Mole's Veggie Adventures app and its effectiveness in increasing FV acceptance among Finnish and Polish preschoolers aged 3 to 6 years.

Methods: A multiprofessional team created an app to be used in ECEC centers in groups of 3 to 10 children. The app aimed to increase vegetable acceptance, and it was built using elements that support the development of self-regulation and social skills. Altogether, 7 Finnish and 4 Polish ECEC centers participated in the study. Before randomization, parents reported background factors and their children's willingness to taste different FVs. The ECEC professionals in the intervention arm were instructed to use the app at least once a week during the 3- to 4-week intervention period. The main outcomes in this unblinded, cluster-randomized study were FV acceptance and relative FV acceptance. The first was calculated as a sum variable describing the children's willingness to taste 25 different FVs, the second as FV acceptance divided by the number of FVs served. We used analysis of covariance to compare the FV acceptance and relative FV acceptance scores between the intervention and control groups at follow-up.

Results: A total of 221 children were included in the analysis. At follow-up, the intervention group (115/221, 52%) had higher FV acceptance scores (baseline adjusted difference of mean 7.22; 95% CI 1.41-13.03) than the control group (106/221, 48%). The intervention effect was parallel for relative FV acceptance scores (baseline adjusted difference of mean 0.28; 95% CI 0.05-0.52).

Conclusions: The Mole's Veggie Adventures app has the potential to increase FV acceptance among preschoolers and can be a valuable tool in supporting food education in ECEC centers. Furthermore, the app can be feasibly incorporated into preschool routines in countries with different educational environments.

Trial Registration: ClinicalTrials.gov NCT05173311; https://tinyurl.com/4vfbh283

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2022;10(1):e30352) doi:10.2196/30352

KEYWORDS

gamification; intervention; behavior change; health game; games for health; smartphone app; mobile phone

Introduction

Background

Most European children do not consume the recommended amount of fruit and vegetables (FVs) [1,2]. Among European countries, Poland and Finland face the same challenges. For instance, there seems to be a large proportion of both Polish and Finnish preadolescents who do not eat FVs daily [3]. Studies among Polish preschoolers are scarce, but in a 2011 report, the proportion of Finnish 6- to 8-year-olds consuming the recommended amount of FVs was less than 5% [4]. More recent studies have observed average consumption to be closer to recommendations, but the consumption of vegetables seems to be lower than that of fruit [5,6]. Indeed, children tend to prefer sweet tastes, as observed in fruits, compared with bitter-tasting foods, such as vegetables [7], and need more taste exposures to accept new vegetables [8]. However, as reassuring evidence suggests that repeated taste exposure and even exposure to picture books can help children to learn to enjoy vegetables [7-9], early childhood is a significant phase to support the formation of healthy eating habits among children.

In both Poland and Finland, most 3- to 6-year-olds attend early childhood education and care centers (ECECs) [10,11]. The general aim of the Finnish curriculum for ECEC is to strengthen skills related to children's well-being [12], such as self-regulation skills, which refer to the ability to monitor and manage emotions and behaviors. Reinforcing self-regulation skills in childhood is important because they are associated with health outcomes later in life [13-15]. Moreover, self-regulation skills are also linked to health behaviors because, for instance, eating is regulated according to internal cues of hunger and fullness [16]. Both the Polish and Finnish recommendations for ECEC encourage ECEC professionals to support the development of self-regulation in eating and to promote a positive attitude toward food and eating [17,18]. Hence, food education is part of the pedagogically guided activities and holistic learning about well-being. However, ECEC professionals lack concrete, age-appropriate, effective, yet appealing tools for food education.

Objective

Mobile devices, such as smartphones and tablets, are ubiquitous and increasingly used at ECEC centers [19]; thus, food education can be delivered using technology. As ECEC centers should provide children with equal possibilities to familiarize themselves with technology and to practice responsible use of digital devices [12], the use of technology per se is beneficial. Furthermore, video games can trigger feelings of joy, intense participation, social interaction, and pleasure [20,21], and their educational use is considered promising [22]. Vast numbers of educational games and apps are already available through the digital distribution platforms Google Play and the App Store, and some of these even focus on food education. Studies reporting on the use of digital games and food-related outcomes are scarce and mostly concentrate on negative outcomes, such as an increase in fast food consumption due to advergaming exposure [23,24]. However, evidence from the United States suggests that video games can positively affect FV intake among 9- to 12-year-olds [25-27]. Moreover, a mobile app including vegetable-based activities has been shown to increase liking and consumption of vegetables among 3- to 6-year-old children in the United Kingdom [28]; however, educational games targeting ECEC environments are lacking. To fill this gap in knowledge, this paper describes the development of the Mole's Veggie Adventures app and its effectiveness in increasing FV acceptance among Finnish and Polish preschoolers.

Methods

App Development

As part of the European Institution of Innovation & Technology (EIT) Food School Network project and together with a software development company specialized in designing, developing, and implementing serious games and gamified solutions (NordicEdu Oy), we designed and pilot-tested the Mole's Veggie Adventures mobile app to increase vegetable acceptance among preschoolers. The University of Helsinki team was in charge of the app development, and the educational content of the app was designed by experts in nutrition science, food education, and ECEC. The content was first created in Finnish and later adapted and translated into English and Polish. The design of the app is described in detail in Multimedia Appendix 1. Briefly, the app was designed to be used in ECEC centers in groups of 3-10 children, but the ECEC professionals were encouraged to adjust the contents to fit the current situation in their group. The primary purpose of the app is to familiarize children with FV and increase FV acceptance. Unlike traditional mobile apps, Mole's Veggie Adventures was built using elements that support the development of self-regulation and social skills. The app consists of 4 seasons, each of which includes 6 FVs. At the time of the intervention, the app listed 6 tasks for each of the vegetables and fruits: (1) Learn, (2) Color, (3) Shape, (4) Taste, (5) Pretend, and (6) Play, and the current version was numbered 0.4.5.0 (7b57516). An updated version of the Mole's Veggie Adventures app is free of charge and available for download in the App Store and Google Play, and Multimedia Appendix 2 provides an overview of the most important sections of the app. All changes were made after the intervention. Table 1 summarizes the main characteristics of the app.

XSL•FO
RenderX

Table 1. Short description (descriptive table modified from the original form [29]) of the Mole's Veggie Adventures app.

Characteristics	Description
Health topics covered	Food behavior (especially acceptability and consumption of FVs[a]); food education
Targeted age group and environment	3- to 6-year-olds; preschool groups
Short description of the game idea	The main character in the game is the Mole, who moves around in a vegetable patch. The game is divided into seasons, each of which contains 6 FVs. The children can familiarize themselves with the FVs by completing different tasks in a group. For each FV, there are adult-led tasks to be completed in groups. In addition, the game includes a Taste Bank, which can be used to record the number of FVs tasted by the group, and Mini-Games, which can be played individually or in pairs. Multimedia Appendices 1 and 2 describe the contents of the game in more detail
Target players	• Individual • Dyad • Small group
Guiding knowledge or behavior change theories, models, or conceptual frameworks	Interactive tailoring (the ECEC[b] professionals can adjust the tasks to be suitable for their group); role-playing (the players can learn from each other and from the ECEC professionals); goal setting and social cognitive theory (the group can decide to taste new vegetables together); learning through play (social interaction, motor skills, self-regulation etc.)
Intended health behavior change	Increase in FV acceptance
Clinical or parental support needed?	ECEC professionals or parents needed in the adult-led sections; mini-games can be played without adults
Data shared with parent or clinician?	No
Type of game	• Active • Role-playing • Educational
Game platforms needed to play the game	• Smartphone • Tablet
Recommended play time	30 min at a time; minimum of 1-2 times a week

[a]FV: fruit and vegetable.

[b]ECEC: early childhood education and care.

Recruitment

To test the effectiveness of the app, we conducted a feasibility study in 2 countries, Finland and Poland, between September and November 2019. On the basis of the literature regarding pilot trial sample size estimation [30], we aimed to recruit 100 children from both countries. Owing to differences in the ECEC systems and cultural environment, we describe the recruitment separately for the 2 countries. In Helsinki, Finland, 12 ECEC center directors were contacted and asked to participate in the study. Of these, 33% (4/12) declined (2 ECEC centers had a busy schedule, 1 did not have enough resources to participate, and in 1 ECEC center, the ECEC professionals were not enthusiastic about the study). In addition, one director could not be reached by email or phone. Thus, 14 groups from 7 public ECEC centers (58% of those invited) agreed to participate in the study. From the consenting groups, we invited all children to participate in the study. Informed consent was requested from legal guardians (later referred to as parents) via the ECEC groups, and the parents of 56% (130/232) of children invited provided their consent to participate in the study. The study was approved by the Education Division of the City of Helsinki, and the University of Helsinki Ethical Review Board in Humanities and Social and Behavioral Sciences deemed the study to be ethically acceptable (Statement 35/2019).

In Poland, the heads of 4 ECEC centers agreed to participate in the study. The study was carried out in 1 public ECEC center in the countryside (Wilczyn) and 2 public (Międzylesie-Warsaw, Piaseczno-Warsaw) and 1 private (Kobyłka-Warsaw) ECEC center in large urban agglomerations. The University of Warsaw's research team organized meetings for parents in each ECEC center. The aims of the meetings were to introduce the goals of the study and to present the educational content of the intervention to the ECEC professionals and parents of the participating children. The parents received informed consent forms in the meetings, and of 213 who were invited, the parents of 196 (92%) children provided their consent to participate in the study. The study procedure was evaluated and approved by the Ethics Committee of the Faculty of Psychology at the University of Warsaw.

Background Characteristics

At baseline, the parents of the participating children completed questionnaires regarding background factors and their children's FV acceptance. They reported the child's gender and birthdate as well as the number of children living in the same household. The number of children living in the same household was categorized into 3 groups: 1 child, 2 children, and 3 or more children. In addition, the parents indicated whether the child had any vegetable- or fruit-related food allergies. The parents reported their highest educational level using 6 predefined

XSL•FO
RenderX

response options (comprehensive school, upper secondary school, vocational school, bachelor's degree, master's degree, and licentiate or doctoral degree), which were categorized into 3 groups: low (comprehensive, upper secondary, or vocational school), middle (bachelor's degree), and high (master's degree or higher).

Outcomes

The parents filled in a questionnaire listing 25 vegetables and fruits and inquiring whether these had been offered to the child during the past 4 weeks and how the child reacted to those that had been served. All the listed vegetables and fruits were introduced in the app. The answer options were 0=*was not offered during the past four weeks*, 1=*refused to touch food*, 2=*touched food but did not put in/near mouth*, 3=*put food to lips but not in mouth*, 4=*put food in mouth but spat out/did not eat*, and 5=*ate food*. A similar questionnaire was used earlier in a UK study examining toddlers' willingness to taste different foods [8]. For each participant, we calculated an FV acceptance score by summing the answers to each of the 25 vegetable and fruit items, with higher scores indicating a higher FV acceptance (theoretical range 0-125). We also calculated the number of FVs served during the past 4 weeks (range 0-25) and used this information to create a relative FV acceptance score (range 0-5) by dividing the FV acceptance score by the number of FVs served.

Intervention and Control Arms

After the parents of the participating children had filled in the baseline questionnaires, the participating ECEC centers (in Finland) or groups within the ECEC centers (in Poland) were randomly allocated into intervention and control arms. In Finland, 4 ECEC centers with 7 groups were randomized into the intervention arm, whereas 7 groups from 3 ECEC centers were in the control arm. In Poland, groups within each ECEC center were evenly randomized into the intervention (5 groups from 4 ECEC centers) and control arms (5 groups from 4 ECEC centers). The study was not blinded. Researchers visited the intervention arm groups and introduced the app to the ECEC professionals. The ECEC professionals received a printed guide, which contained instructions and information about the app, and a PDF version of the guide was also available through the app. The ECEC professionals were instructed to use the app with a tablet computer at least one to two times a week during the intervention period (3-4 weeks) and to record the number of tasks completed by their group in a logbook. In addition, we recommended that each group focus on at least six vegetables or fruits during the intervention period. The ECEC professionals also provided quantitative and qualitative feedback using a feedback form. Written feedback was used to update the app after the study period. The control arm groups were instructed to continue their normal routines during the intervention period. They were instructed to refrain from introducing any novel food education methods during the intervention period. After the intervention period, follow-up questionnaires inquiring about the children's FV acceptance were distributed to the parents of the participating children. To treat the intervention and control groups democratically, the app was introduced to the control ECECs after the study period. The trial was not registered because the study did not assess health outcomes and was thus not a clinical trial (World Health Organization defines a clinical trial as "any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes" [31]).

Statistical Analysis

We used an analysis of covariance-type linear model to compare the FV acceptance and relative FV acceptance scores at follow-up. These models were adjusted with baseline FV acceptance score categories (*missing*, *lower than median*, and *median or higher*). To investigate the sensitivity of the results to the choice of the number of baseline categories, we also used baseline FV acceptance scores in tenths (unadjusted FV acceptance score) and sevenths (relative FV acceptance score) in the models.

Results

Altogether, 67.8% (221/326) of children had data on FV acceptance and relative FV acceptance scores at follow-up and were included in the analyses. The participating children were on average aged 5.0 years (SD 1.2 years). About half of the participants were girls, and slightly more children participated from the Polish than from the Finnish ECEC centers (Table 2). In 55.7% (123/221) of the participating families, at least one of the parents had a master's degree or higher education, and most of the respondents (the person who filled in the questionnaires on behalf of the child) were mothers.

Table 3 shows the FV acceptance and relative FV acceptance scores at baseline and at follow-up. At follow-up, the FV acceptance score was 78.5 in the intervention group and 72.4 in the control group, whereas the values for relative FV acceptance scores were 3.97 and 3.75, respectively (Table 3). A score of approximately 3 means that on average, the children put the FVs on the lips but not in the mouth, whereas a score of approximately 4 implies that, on average, the children put the FVs in the mouth but did not eat them. When adjusted for baseline FV acceptance score category, participants in the intervention arm scored higher than control participants (Δ estimate 7.22; 95% CI 1.41-13.03). This corresponds to approximately a 10% improvement because of the intervention compared with a control group participant with the same baseline score. Similarly, relative FV acceptance scores at follow-up were, on average, 0.28 higher (+7%) in the intervention arm group than in the control arm group (95% CI 0.05-0.52). Regarding FV acceptance scores, the sensitivity analyses (data not shown) showed a similar and consistently significant intervention effect (Δ estimate 6.38; 95% CI 0.69-12.07), whereas a smaller and borderline significant intervention effect (Δ estimate 0.19; 95% CI −0.03-0.41) was detected for the relative FV acceptance score.

On average, the intervention arm groups used the app 1.9 times/week during the intervention period. The frequency of app use was missing from one group, but the number of completed tasks was as recommended, suggesting sufficient compliance. Furthermore, 17% (2/12) of groups did not use the app as instructed, that is, they used the app less than once a

XSL•FO
RenderX

week and did not complete tasks related to at least six FVs. Both groups were from Finnish ECEC centers. The app was typically used in a group of 2-10 children in the Finnish ECEC centers, whereas the usual group size in the Polish ECEC centers was 24-25 children. The most popular FVs chosen by the intervention arm groups were blueberries (11/12, 92% of the groups completed related tasks); lettuce (10/12, 83%); mushrooms (9/12, 75%); kidney, brown, and black beans (9/12, 75%); beetroot (8/12, 67%); and squash (8/12, 67%).

Table 2. Description of the study population (N=221).

Characteristics	Total (N=221), n (%)	Intervention (n=115), n (%)	Control (n=106), n (%)
Gender			
Girls	120 (54.3)	67 (58.3)	53 (50.0)
Boys	100 (45.2)	47 (40.9)	53 (50.0)
Missing	1 (0.5)	1 (0.9)	0 (0.0)
Country			
Finland	95 (43.0)	50 (43.5)	45 (42.5)
Poland	126 (57.0)	65 (56.5)	61 (57.5)
Vegetable or fruit allergy			
No	208 (94.1)	108 (93.9)	100 (94.3)
Yes	12 (5.4)	7 (6.1)	5 (4.7)
Missing	1 (0.5)	0 (0.0)	1 (0.9)
Number of children living in the same household			
One	56 (25.3)	27 (23.5)	29 (27.4)
Two	124 (56.1)	62 (53.9)	62 (58.5)
Three or more	36 (16.3)	22 (19.1)	14 (13.2)
Missing	5 (2.3)	4 (3.5)	1 (0.9)
Respondent			
Father	27 (12.2)	15 (13.0)	12 (11.3)
Mother	193 (87.3)	100 (87.0)	93 (87.7)
Missing	1 (0.5)	0 (0.0)	1 (0.9)
Parental educational level			
Upper secondary school or lower	55 (24.9)	32 (27.8)	23 (21.7)
Bachelor's degree or equivalent	39 (17.6)	21 (18.3)	18 (17.0)
Master's degree or higher	123 (55.7)	60 (52.2)	63 (59.4)
Missing	4 (1.8)	2 (1.7)	2 (1.9)

Table 3. FV[a] acceptance and relative FV acceptance scores in the intervention (n=82-115) and control (n=79-106) groups at baseline and at follow-up.

Characteristics	Baseline, mean (SD)	Follow-up, mean (SD)
FV acceptance score[b]		
Intervention group	70.6 (25.5)	78.5 (30.6)
Control group	70.2 (25.0)	72.4 (26.2)
Relative FV acceptance score[c]		
Intervention group	3.84 (1.06)	3.97 (1.03)
Control group	3.77 (1.10)	3.75 (1.01)

[a]FV: fruit and vegetable.

[b]FV acceptance score: sum variable describing willingness to taste the 25 FVs listed; higher score indicates higher FV acceptance (theoretical range 0-125).

[c]FV acceptance score: FV acceptance score divided by the number of FVs served (range 0-5).

XSL•FO
RenderX

Discussion

Principal Findings

This paper describes the design and pilot-testing of the Mole's Veggie Adventures app, which aimed to increase FV acceptance among 3- to 6-year-old preschoolers in Finland and Poland. Our pilot study showed a favorable and meaningful intervention effect; compared with the control arm participants, the participants in the intervention arm had higher FV acceptance scores at follow-up 3-4 weeks after baseline. Thus, the app can be considered an effective food education tool in an early education environment. Earlier studies have shown that video games are effective in increasing FV consumption among 9- to 12-year-old children [25-27], and promising results among preschool aged participants have also been obtained [28]. To the best of our knowledge, the current food education tool is the first to be used routinely in a preschool environment. As healthy food behaviors, such as frequent and diverse FV consumption, are typically adopted in childhood and may track into adulthood [32-34], early childhood is a crucial time to intervene.

Comparison With Earlier Work

Some serious games emphasize increased knowledge [35], whereas others aim to incorporate multiple theory-driven behavior change techniques, such as tailoring, goal setting, problem solving, and feedback, into a fun and attractive game [36]. In Mole's Veggie Adventures, no specific behavior change technique was selected, but several of them were included in the game mechanics. For instance, the ECEC professionals were instructed to select those FVs for discussion that they deemed most important for their group. Moreover, the game includes Mini-Games with educational and knowledge-enhancing content. In addition, advergame researchers have described multiple methods that can potentially influence player behavior [37]. Advertising in games can appear at different levels of the game and in many forms, for example, as product placement, background presentation, and engagement via interactivity. In the case of the Mole's Veggie Adventures app, all the aforementioned expositions of FVs are present, as the product, FVs, appears both in the background and is subject to manipulation itself, potentially enhancing FV acceptance among children. In addition, because of the various types of tasks and activities in the game, the children do not familiarize themselves with FVs only virtually but also in reality. The diverse stimuli with FVs as the main characters may encourage children to become familiar with them. However, the extent to which such an intervention could realistically affect behavior warrants further research.

The Mole's Veggie Adventures app includes a strong social aspect. It has been shown that ECEC professionals' opinions may contribute to children's food consumption [38]. In addition, peers may also act as role models for preschoolers [39]. As our game was used in the ECEC centers in a group of preschoolers, it may have offered opportunities to model—for better or worse—the early educators as well as other children and motivated children to try new FVs. Social interaction provides opportunities for problem solving and peer engagement, which in turn, can cultivate useful skills such as negotiation and cooperation [40]. In addition, approval from the early educator, supporting comments from the group, and the opportunity to boast and present the results of tasks in the group may have been rewarding. Previous serious game research has also suggested that engaging parents—gatekeepers of the home environment [41]—may be critical in changing child behavior [26,27]. Bearing this in mind, we updated the Mole's Veggie Adventures app after the intervention to better fit both the preschool and home environments (Multimedia Appendices 1 and 2). We also encourage future game designers to consider including a parental component to ensure adult support in all environments relevant to the child.

Fun is an essential part of playing games and can produce intrinsic motivation in players [42]. However, it remains unknown how certain target groups (eg, preschoolers) comprehend and experience fun or how to use fun to design games to bring about larger or more consistent changes in health behaviors [43]. Baranowski et al [42] contemplated the building blocks of fun and suggested that fun in games is probably a combination of interaction, overcoming challenges, making choices, and detecting their consequences without risking oneself, receiving feedback, increasing difficulty through levels of game play, and using personally relevant stories and characteristics in meaningful situations. To ensure that the Mole's Veggie Adventures app would be perceived as fun, preschoolers participated in the development process (see Multimedia Appendix 1 for details). In addition, the game incorporated elements, such as physical play, invented stories, and adult-led activities, which have been identified as occasions for fun and shared humor among preschoolers [44]. A growing consensus describes play as an intrinsically motivated activity that results in joyful discovery [40]; thus, it is presumable that the app, by covering various forms of play (ie, active physical and pretend play), was indeed perceived as fun by the preschoolers.

Although games can deliver food education in an enticing way, not all behaviors encouraged by games are beneficial. Excessive gaming can evoke negative psychosocial effects [20] and even cause addictive behaviors [45]. Possible adverse effects include increased impulsivity [46], and as impulse control is an element in self-regulation, games can impair the development of self-regulation skills. Poor self-regulation skills in childhood have been linked to diminished social and cognitive outcomes later in life [47] and may also be associated with adverse health outcomes such as overweight and obesity or increased screen time [48-52]. To avoid these pitfalls, the Mole's Veggie Adventures app was designed to support the development of self-regulation skills. For instance, the app includes elements that require peaceful action and waiting. Moreover, the app is mostly intended to be used with an ECEC professional, whose role as a coregulator is significant in strengthening self-regulation skills [53]. Subdued, mild colors and delicate music allow the child to focus on the educational content and could potentially prevent impulsivity.

XSL•FO
RenderX

Study Strengths and Limitations

The strengths of the study include testing in 2 countries, Finland and Poland. The 2-country setting allowed us to recruit a larger sample, which in turn, enabled the detection of the intervention effect. Another strength is the random allocation of ECEC centers into the intervention and control arms. Therefore, it is unlikely that the outcome was confounded by uncontrolled variables. Moreover, we examined one specific outcome (FV acceptance) instead of testing for multiple outcomes, which could have led to type 1 error [54]. The app development process was extensive and included cocreation with the target group and a multidisciplinary research team as well as prepiloting of the demo version in Finnish preschools (see Multimedia Appendix 1 for details). Most intervention arm groups used the app as instructed, suggesting moderate feasibility.

Although the study was able to demonstrate favorable changes in FV acceptance, it also had some limitations that should be addressed. First, our study was not blinded, and the participating early educators in the intervention group knew that the children's FV acceptance was being measured. The parents of the participating children were also aware of the intervention. However, the app was used in the ECEC centers, whereas parents reported FV acceptance, and thus, the parents did not know exactly how much their children had used the app. In Finland, randomization was conducted at the preschool level to avoid contamination. Owing to nonexistent between-group communication among parents, contamination was not considered probable in Poland. Second, only 72.9% (161/221) of participants had data at both baseline and follow-up. To use data from as many participants as possible, we categorized participants into 3 groups based on their baseline data: missing, lower than median, and median or higher. To determine how the categorization affected the results, we ran multiple sensitivity analyses, which yielded parallel results. Third, as the app was used in a group, we did not know which individual children in the intervention arm groups participated in the game sessions. In addition, as the degree of implementation varied between the ECEC groups, this could have attenuated the observed effects. Fourth, our sample was relatively highly educated, and thus, the results may not be generalizable to socioeconomically disadvantaged groups. Owing to differences in cultural environment, the recruitment process was carried out differently in the 2 countries, which probably resulted in differing participation rates (56% in Finland vs 92% in Poland). Thus, the Polish sample might have been more representative of the target population than the Finnish sample. Furthermore, because of the limited time frame set by the funding period, the intervention was relatively short. As children need repeated taste exposures, preferably integrated with sensory learning strategies as well as nutrition education to get used to different vegetables [55], it is possible that a longer intervention period would have been needed to achieve more prominent and permanent results. However, we realistically aimed to increase FV acceptability, not FV consumption, which would probably require more time. Future studies should include postintervention follow-up to examine the stability of the intervention effects.

Conclusions

In summary, the Mole's Veggie Adventures app has the potential to increase FV acceptance among preschoolers. The app can support food education and be incorporated into the preschool curriculum in countries with different educational environments, such as Finland and Poland. When designing serious games for preschoolers, game designers should consider including both home- and preschool-based components to ensure adult endorsement in all relevant environments, which could result in even stronger effects. Future studies should aim to identify the game mechanisms that best support children in making behavior changes.

Acknowledgments

The authors thank the participating children, families, and municipalities, the early childhood education and care centers, and the early childhood education and care professionals. The authors are also grateful to Dr Elena Santa Cruz, Dr Alan Roberts, Professor Kate Harvey, and Professor Tom Baranowski for their collaboration in this project. This work was funded by EIT Food (*The EIT Food School Network: Integrating solutions to improve eating habits and reduce food wastage* # 18145, # 19057, # 20129). EIT Food is the Innovation Community on Food of the EIT, a body of the European Union, under Horizon 2020, the EU Framework Programme for Research and Innovation. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Open access was funded by the Helsinki University Library.

Authors' Contributions

HV, E Skaffari, ML, CR, E Suhonen, NS, and ME participated in the development process of the app; KW and JB translated the app into Polish and adapted it to the Polish cultural context; HV, E Skaffari, KW, JB, and ME designed the study; JB and ME were responsible for funding acquisition; ME was responsible for leadership in research activity planning and execution; HV and KW managed and coordinated research activities; HV, E Skaffari, KW, JB, SK, RM, and MH collected the data and participated in data curation; HV, JN, and ME designed the statistical analyses; HV analyzed the data and wrote the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

The authors are the developers of the Mole's Veggie Adventures mobile app. The authors have no other relationships or activities that could potentially be construed as a conflict of interest with the present work.

XSL•FO
RenderX

Editorial Notice

This randomized study was only retrospectively registered. The authors explained that their study "did not assess health outcomes and was not a clinical trial". However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Detailed description of the application development process.
[DOCX File , 175 KB - mhealth_v10i1e30352_app1.docx]

Multimedia Appendix 2
Video demonstrating the most important sections of the current version of the Mole's Veggie Adventures application. Note that the version used in the current study differed slightly from the current version.
[MP4 File (MP4 Video), 31691 KB - mhealth_v10i1e30352_app2.mp4]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2541 KB - mhealth_v10i1e30352_app3.pdf]

References

1. Lynch C, Kristjansdottir AG, Te Velde SJ, Lien N, Roos E, Thorsdottir I, et al. Fruit and vegetable consumption in a sample of 11-year-old children in ten European countries--the PRO GREENS cross-sectional survey. Public Health Nutr 2014 Nov;17(11):2436-2444. [doi: 10.1017/S1368980014001347] [Medline: 25023091]
2. Cardon G, De Bourdeaudhuij I, Iotova V, Latomme J, Socha P, Koletzko B, ToyBox-study group. Health related behaviours in normal weight and overweight preschoolers of a large pan-european sample: the ToyBox-study. PLoS One 2016;11(3):e0150580 [FREE Full text] [doi: 10.1371/journal.pone.0150580] [Medline: 26950063]
3. Inchley J, Currie D, Budisavljevic S, Torsheim T, Jåstad A, Cosma A, et al. Spotlight on adolescent health and well-being: findings from the 2017/2018 Health Behaviour in School-aged Children (HBSC) survey in Europe and Canada. WHO Regional Office for Europe. 2020. URL: https://apps.who.int/iris/bitstream/handle/10665/332091/9789289055000-eng.pdf [accessed 2021-05-07]
4. Eloranta AM, Lindi V, Schwab U, Kiiskinen S, Kalinkin M, Lakka HM, et al. Dietary factors and their associations with socioeconomic background in Finnish girls and boys 6-8 years of age: the PANIC Study. Eur J Clin Nutr 2011 Nov;65(11):1211-1218. [doi: 10.1038/ejcn.2011.113] [Medline: 21697818]
5. Hauta-Alus HH, Korkalo L, Holmlund-Suila EM, Rosendahl J, Valkama SM, Enlund-Cerullo M, et al. Food and nutrient intake and nutrient sources in 1-year-old infants in Finland: a cross-sectional analysis. Nutrients 2017 Dec 01;9(12):1309 [FREE Full text] [doi: 10.3390/nu9121309] [Medline: 29194422]
6. Korkalo L, Nissinen K, Skaffari E, Vepsäläinen H, Lehto R, Kaukonen R, et al. The contribution of preschool meals to the diet of Finnish preschoolers. Nutrients 2019 Jul 05;11(7):1531 [FREE Full text] [doi: 10.3390/nu11071531] [Medline: 31284433]
7. Forestell CA. Flavor perception and preference development in human infants. Ann Nutr Metab 2017;70 Suppl 3:17-25 [FREE Full text] [doi: 10.1159/000478759] [Medline: 28903110]
8. Houston-Price C, Owen L, Kennedy O, Hill C. Parents' experiences of introducing toddlers to fruits and vegetables through repeated exposure, with and without prior visual familiarization to foods: evidence from daily diaries. Food Qual Prefer 2019 Jan;71:291-300 [FREE Full text] [doi: 10.1016/j.foodqual.2018.08.003]
9. Owen LH, Kennedy OB, Hill C, Houston-Price C. Peas, please! Food familiarization through picture books helps parents introduce vegetables into preschoolers' diets. Appetite 2018 Sep 01;128:32-43. [doi: 10.1016/j.appet.2018.05.140] [Medline: 29807124]
10. Säkkinen S, Kuoppala T. Varhaiskasvatus 2019. Finnish Institute for Health and Welfare. 2020. URL: https://www.julkari.fi/bitstream/handle/10024/140541/Tr33_20.pdf?sequence=5&isAllowed=y [accessed 2021-05-07]
11. Share of children covered by preschool education. Statistics Poland. 2021. URL: https://bdl.stat.gov.pl/BDL/dane/podgrup/tablica [accessed 2021-04-09]
12. National core curriculum for early childhood education and care 2018. Finnish National Agency for Education. 2019. URL: https://www.oph.fi/sites/default/files/documents/varhaiskasvatussuunnitelman_perusteet.pdf [accessed 2021-05-07]
13. Seeyave DM, Coleman S, Appugliese D, Corwyn RF, Bradley RH, Davidson NS, et al. Ability to delay gratification at age 4 years and risk of overweight at age 11 years. Arch Pediatr Adolesc Med 2009 Apr;163(4):303-308 [FREE Full text] [doi: 10.1001/archpediatrics.2009.12] [Medline: 19349558]
14. Montroy JJ, Bowles RP, Skibbe LE, McClelland MM, Morrison FJ. The development of self-regulation across early childhood. Dev Psychol 2016 Nov;52(11):1744-1762 [FREE Full text] [doi: 10.1037/dev0000159] [Medline: 27709999]

XSL•FO
RenderX

15. Anderson SE, Sacker A, Whitaker RC, Kelly Y. Self-regulation and household routines at age three and obesity at age eleven: longitudinal analysis of the UK Millennium Cohort Study. Int J Obes (Lond) 2017 Dec;41(10):1459-1466 [FREE Full text] [doi: 10.1038/ijo.2017.94] [Medline: 28435162]
16. Herman C, Polivy J. The self-regulation of eating: theoretical and practical problems. In: Baumeister RF, Vohs KD, editors. Handbook of Self-Regulation: Research, Theory, and Applications 1st Edition. New York: The Guilford Press; 2004:1-574.
17. National Nutrition Council. Health and joy from food - meal recommendations for early childhood education and care. National Nutrition Council, Finnish National Agency for Education and National Institute for Health and Welfare. 2018. URL: https://www.julkari.fi/bitstream/handle/10024/135969/URN_ISBN_978-952-343-033-4.pdf?sequence=1&isAllowed=y [accessed 2021-05-07]
18. Komentarzem Z. Podstawa programowa wychowania przedszkolnego i ksztalcenia ogólnego dla szkoly podstawowej. Osrodek Rozwoju Edukacji. 2018. URL: https://www.ore.edu.pl/wp-content/uploads/2018/03/podstawa-programowa-wychowania-przedszkolnego-i-ksztalcenia-ogolnego-dla-szkoly-podstawowej-z-komentarzem.pdf [accessed 2021-05-07]
19. Koivula M, Mustola M. Varhaiskasvatuksen digiloikka ja muuttuva sukupolvijärjestys? Jännitteitä lastentarhanopettajien ja lasten kohtaamisissa digitaalisen teknologian äärellä. Kasvatus & Aika 2017;11(3):37-50 [FREE Full text]
20. Mitchell A, Savill-Smith C. The use of computer and video games for learning - a review of the literature. Learning and Skills Development Agency. 2004. URL: https://dera.ioe.ac.uk/5270/7/041529_Redacted.pdf [accessed 2021-05-07]
21. Vorderer P, Hartmann T, Klimmt C. Explaining the enjoyment of playing video games: the role of competition. In: Proceedings of the Second International Conference on Entertainment Computing. 2003 Presented at: Second International Conference on Entertainment Computing; May 8 - 10, 2003; Pittsburgh Pennsylvania USA p. 1-9 URL: https://dl.acm.org/doi/abs/10.5555/958720.958735?preflayout=tabs
22. de Freitas S. Are games effective learning tools? A review of educational games. Edu Technol Soc 2018;21(2):74-84 [FREE Full text]
23. Montgomery K, Grier S, Chester J, Dorfman L. Food marketing in the digital age: a conceptual framework and agenda for research. Center for Digital Democracy. 2011. URL: https://www.democraticmedia.org/sites/default/files/Digital_Food_Mktg_Conceptual_Model%20Report.pdf [accessed 2021-05-07]
24. Nairn A, Haiming H. Advergames: it's not child's play - a review of research. Institute for Policy Research, University of BATH. 2012. URL: https://www.bath.ac.uk/publications/advergames-its-not-childs-play/attachments/ipr-policy-brief-advergames-its-not-childs-play.pdf [accessed 2021-05-07]
25. Baranowski T, Baranowski J, Thompson D, Buday R, Jago R, Griffith MJ, et al. Video game play, child diet, and physical activity behavior change a randomized clinical trial. Am J Prev Med 2011 Jan;40(1):33-38 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.029] [Medline: 21146765]
26. Cullen KW, Liu Y, Thompson DI. Meal-specific dietary changes from squires quest! II: a serious video game intervention. J Nutr Educ Behav 2016 May;48(5):326-330 [FREE Full text] [doi: 10.1016/j.jneb.2016.02.004] [Medline: 27169641]
27. Thompson D, Bhatt R, Vazquez I, Cullen KW, Baranowski J, Baranowski T, et al. Creating action plans in a serious video game increases and maintains child fruit-vegetable intake: a randomized controlled trial. Int J Behav Nutr Phys Act 2015;12:39 [FREE Full text] [doi: 10.1186/s12966-015-0199-z] [Medline: 25890060]
28. Farrow C, Belcher E, Coulthard H, Thomas JM, Lumsden J, Hakobyan L, et al. Using repeated visual exposure, rewards and modelling in a mobile application to increase vegetable acceptance in children. Appetite 2019 Oct 01;141:104327. [doi: 10.1016/j.appet.2019.104327] [Medline: 31228505]
29. Baranowski T. Descriptions for articles introducing a new game for health. Games Health J 2014 Apr;3(2):55-56. [doi: 10.1089/g4h.2014.0008] [Medline: 26196042]
30. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
31. International Clinical Trials Registry Platform (ICTRP). World Health Organization. URL: https://www.who.int/clinical-trials-registry-platform [accessed 2021-05-07]
32. Kelder SH, Perry CL, Klepp KI, Lytle LL. Longitudinal tracking of adolescent smoking, physical activity, and food choice behaviors. Am J Public Health 1994 Jul;84(7):1121-1126. [doi: 10.2105/ajph.84.7.1121] [Medline: 8017536]
33. Lien N, Lytle LA, Klepp KI. Stability in consumption of fruit, vegetables, and sugary foods in a cohort from age 14 to age 21. Prev Med 2001 Sep;33(3):217-226. [doi: 10.1006/pmed.2001.0874] [Medline: 11522162]
34. Mikkilä V, Räsänen L, Raitakari OT, Pietinen P, Viikari J. Consistent dietary patterns identified from childhood to adulthood: the cardiovascular risk in Young Finns Study. Br J Nutr 2005 Jun;93(6):923-931. [doi: 10.1079/bjn20051418] [Medline: 16022763]
35. Peng W. Design and evaluation of a computer game to promote a healthy diet for young adults. Health Commun 2009 Mar;24(2):115-127. [doi: 10.1080/10410230802676490] [Medline: 19280455]
36. Thompson D, Baranowski T, Buday R, Baranowski J, Thompson V, Jago R, et al. Serious video games for health how behavioral science guided the development of a serious video game. Simul Gaming 2010 Aug 1;41(4):587-606 [FREE Full text] [doi: 10.1177/1046878108328087] [Medline: 20711522]

XSL•FO
RenderX

37. Tina W, Buckner K. Receptiveness of gamers to embedded brand messages in advergames. J Interact Advert 2006 Sep;7(1):3-32. [doi: 10.1080/15252019.2006.10722123]
38. Lehto R, Ray C, Vepsäläinen H, Korkalo L, Nissinen K, Skaffari E, et al. Early educators' practices and opinions in relation to pre-schoolers' dietary intake at pre-school: case Finland. Public Health Nutr 2019 Jun;22(9):1567-1575. [doi: 10.1017/S1368980019000077] [Medline: 30782234]
39. Ward SA, Bélanger MF, Donovan D, Carrier N. Relationship between eating behaviors and physical activity of preschoolers and their peers: a systematic review. Int J Behav Nutr Phys Act 2016 Apr 14;13:50 [FREE Full text] [doi: 10.1186/s12966-016-0374-x] [Medline: 27075482]
40. Yogman M, Garner A, Hutchinson J, Hirsh-Pasek K, Golinkoff RM, Committee on Psychosocial Aspects of Child and Family Health, Council on Communications and Media. The power of play: a pediatric role in enhancing development in young children. Pediatrics 2018 Sep;142(3):e20182058 [FREE Full text] [doi: 10.1542/peds.2018-2058] [Medline: 30126932]
41. Gruber KJ, Haldeman LA. Using the family to combat childhood and adult obesity. Prev Chronic Dis 2009 Jul;6(3):A106 [FREE Full text] [Medline: 19527578]
42. Baranowski T, Thompson D, Buday R, Lu AS, Baranowski J. Design of video games for children's diet and physical activity behavior change. Int J Comput Sci Sport 2010;9(2):3-17 [FREE Full text] [Medline: 25364331]
43. Baranowski T. Games for health research—past, present, and future. Präv Gesundheitsf 2018 Jul 12;13(4):333-336 [FREE Full text] [doi: 10.1007/s11553-018-0657-y]
44. Stenius TH, Karlsson L, Sivenius A. Young children’s humour in play and moments of everyday life in ECEC centres. Scand J Educ Res 2021 Jan 06:1-15. [doi: 10.1080/00313831.2020.1869084]
45. Gentile DA, Bailey K, Bavelier D, Brockmyer JF, Cash H, Coyne SM, et al. Internet gaming disorder in children and adolescents. Pediatrics 2017 Nov;140(Suppl 2):81-85 [FREE Full text] [doi: 10.1542/peds.2016-1758H] [Medline: 29093038]
46. Dong G, Potenza MN. A cognitive-behavioral model of Internet gaming disorder: theoretical underpinnings and clinical implications. J Psychiatr Res 2014 Nov;58:7-11 [FREE Full text] [doi: 10.1016/j.jpsychires.2014.07.005] [Medline: 25062755]
47. Mischel W, Shoda Y, Rodriguez MI. Delay of gratification in children. Science 1989 May 26;244(4907):933-938. [doi: 10.1126/science.2658056] [Medline: 2658056]
48. Francis LA, Susman EJ. Self-regulation and rapid weight gain in children from age 3 to 12 years. Arch Pediatr Adolesc Med 2009 Apr;163(4):297-302. [doi: 10.1001/archpediatrics.2008.579] [Medline: 19349557]
49. Nederkoorn C, Braet C, Van Eijs Y, Tanghe A, Jansen A. Why obese children cannot resist food: the role of impulsivity. Eat Behav 2006 Nov;7(4):315-322. [doi: 10.1016/j.eatbeh.2005.11.005] [Medline: 17056407]
50. Guerrieri R, Nederkoorn C, Jansen A. The interaction between impulsivity and a varied food environment: its influence on food intake and overweight. Int J Obes (Lond) 2008 Apr;32(4):708-714. [doi: 10.1038/sj.ijo.0803770] [Medline: 18059403]
51. Caleza C, Yañez-Vico RM, Mendoza A, Iglesias-Linares A. Childhood Obesity and Delayed Gratification Behavior: A Systematic Review of Experimental Studies. J Pediatr 2016 Feb;169:201-7.e1. [doi: 10.1016/j.jpeds.2015.10.008] [Medline: 26563536]
52. Radesky JS, Silverstein M, Zuckerman B, Christakis DA. Infant self-regulation and early childhood media exposure. Pediatrics 2014 May;133(5):1172-1178 [FREE Full text] [doi: 10.1542/peds.2013-2367] [Medline: 24733868]
53. Suhonen E, Sajaniemi NK, Alijoki A, Nislin MA. Children’s biological givens, stress responses, language and cognitive abilities and family background after entering kindergarten in toddlerhood. Early Child Devel Care 2016 Aug 20;188(3):345-358. [doi: 10.1080/03004430.2016.1218157]
54. Simes RJ. An improved Bonferroni procedure for multiple tests of significance. Biometrika 1986;73(3):751-754. [doi: 10.1093/biomet/73.3.751]
55. Nekitsing C, Hetherington MM, Blundell-Birtill P. Developing healthy food preferences in preschool children through taste exposure, sensory learning, and nutrition education. Curr Obes Rep 2018 Mar;7(1):60-67 [FREE Full text] [doi: 10.1007/s13679-018-0297-8] [Medline: 29446037]

Abbreviations

ECEC: early childhood education and care
EIT: European Institution of Innovation & Technology
FV: fruit and vegetable

XSL•FO
RenderX

Edited by R Kukafka, G Eysenbach; submitted 11.05.21; peer-reviewed by C Farrow, E Haycraft; comments to author 28.06.21; revised version received 08.07.21; accepted 13.10.21; published 04.01.22.

Please cite as:
Vepsäläinen H, Skaffari E, Wojtkowska K, Barlińska J, Kinnunen S, Makkonen R, Heikkilä M, Lehtovirta M, Ray C, Suhonen E, Nevalainen J, Sajaniemi N, Erkkola M
A Mobile App to Increase Fruit and Vegetable Acceptance Among Finnish and Polish Preschoolers: Randomized Trial
JMIR Mhealth Uhealth 2022;10(1):e30352
URL: https://mhealth.jmir.org/2022/1/e30352
doi:10.2196/30352
PMID:34982718

©Henna Vepsäläinen, Essi Skaffari, Katarzyna Wojtkowska, Julia Barlińska, Satu Kinnunen, Riikka Makkonen, Maria Heikkilä, Mikko Lehtovirta, Carola Ray, Eira Suhonen, Jaakko Nevalainen, Nina Sajaniemi, Maijaliisa Erkkola. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 04.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Investigating When, Which, and Why Users Stop Using a Digital Health Intervention to Promote an Active Lifestyle: Secondary Analysis With A Focus on Health Action Process Approach–Based Psychological Determinants

Helene Schroé[1,2], MSc; Geert Crombez[2], PhD; Ilse De Bourdeaudhuij[1], PhD; Delfien Van Dyck[1], PhD

[1]Department of Movement and Sports Sciences, Faculty of Medicine and Health, Ghent University, Ghent, Belgium

[2]Department of Experimental-Clinical and Health Psychology, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium

Corresponding Author:
Helene Schroé, MSc
Department of Movement and Sports Sciences
Faculty of Medicine and Health
Ghent University
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 9264 63 63
Email: helene.schroe@ugent.be

Abstract

Background: Digital health interventions have gained momentum to change health behaviors such as physical activity (PA) and sedentary behavior (SB). Although these interventions show promising results in terms of behavior change, they still suffer from high attrition rates, resulting in a lower potential and accessibility. To reduce attrition rates in the future, there is a need to investigate the reasons why individuals stop using the interventions. Certain demographic variables have already been related to attrition; however, the role of psychological determinants of behavior change as predictors of attrition has not yet been fully explored.

Objective: The aim of this study was to examine when, which, and why users stopped using a digital health intervention. In particular, we aimed to investigate whether psychological determinants of behavior change were predictors for attrition.

Methods: The sample consisted of 473 healthy adults who participated in the intervention MyPlan 2.0 to promote PA or reduce SB. The intervention was developed using the health action process approach (HAPA) model, which describes psychological determinants that guide individuals in changing their behavior. If participants stopped with the intervention, a questionnaire with 8 question concerning attrition was sent by email. To analyze when users stopped using the intervention, descriptive statistics were used per part of the intervention (including pre- and posttest measurements and the 5 website sessions). To analyze which users stopped using the intervention, demographic variables, behavioral status, and HAPA-based psychological determinants at pretest measurement were investigated as potential predictors of attrition using logistic regression models. To analyze why users stopped using the intervention, descriptive statistics of scores to the attrition-related questionnaire were used.

Results: The study demonstrated that 47.9% (227/473) of participants stopped using the intervention, and drop out occurred mainly in the beginning of the intervention. The results seem to indicate that gender and participant scores on the psychological determinants action planning, coping planning, and self-monitoring were predictors of first session, third session, or whole intervention completion. The most endorsed reasons to stop using the intervention were the time-consuming nature of questionnaires (55%), not having time (50%), dissatisfaction with the content of the intervention (41%), technical problems (39%), already meeting the guidelines for PA/SB (31%), and, to a lesser extent, the experience of medical/emotional problems (16%).

Conclusions: This study provides some directions for future studies. To decrease attrition, it will be important to personalize interventions on different levels, questionnaires (either for research purposes or tailoring) should be kept to a minimum especially in the beginning of interventions by, for example, using objective monitoring devices, and technical aspects of digital health interventions should be thoroughly tested in advance.

Trial Registration: ClinicalTrials.gov NCT03274271; https://clinicaltrials.gov/ct2/show/NCT03274271

XSL•FO
RenderX

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-019-3456-7

(JMIR Mhealth Uhealth 2022;10(1):e30583) doi:10.2196/30583

KEYWORDS

digital health; psychosocial determinants; health action process approach; physical activity; sedentary behavior; attrition; dropout; mobile health; healthy life style; health behaviors

Introduction

Digital health interventions have gained momentum to change health behaviors such as physical activity (PA) and sedentary behavior (SB) [1,2]. Their potential value lies in the ability to reach large groups in a personal, cost-effective, and time-efficient way [3-5]. Previous interventions have shown promising results in terms of behavior change [5-9]. Nevertheless, there are differences in use and completion [10-12]. It is important to ensure that participants do not drop out of the intervention. So far, attrition rates in digital health interventions are high (50% to 80%) [10,13,14]. As a result, interventions may lose part of their potential and accessibility. Also, effective evaluation of trials becomes a challenge. In order to reduce attrition, there is a need to investigate the reasons why individuals drop out. This question is often not addressed, as most studies focus on the effectiveness of interventions [15]. The answers may, however, provide valuable information in developing future digital health interventions [10]. In response to this problem, there has been a call for a science of attrition [10].

In order to understand attrition, 3 questions could be considered [10]. First, when do users stop using the intervention? Answers to this question may allow identification of weak parts of an intervention and may help in redesigning, restructuring, or removing certain parts. Most interventions only describe attrition rates at the end of the intervention. However, reporting attrition proportions at several time points can provide valuable information. For example, different patterns of attrition may occur: (1) a constant proportion of users may drop out of the intervention, (2) users may stay in the intervention first out of curiosity, which relates to the novelty effect (ie, the human tendency for heightened engagement to a novel phenomenon [16]), and then drop out when the novelty has worn off and eventually a stable group remains, (3) a group of users drops out of the intervention immediately and a stable group of users remains [10]. Each pattern could indicate different underlying causes of attrition.

Second, which users stop using the intervention? An answer to this question may direct researchers to tailor the content of the intervention to particular subgroups. Demographic variables such as being male [2,17,18], having a young age [17-21], having a lower educational level [22], and not having a partner [17] have been related to higher attrition rates in digital health interventions. The role of BMI in relation to attrition shows inconsistent results [21,23]. Also, the behavioral status of the participant at the start of the intervention may have an effect. Participants meeting the guidelines for moderate physical activity and for vegetable consumption at baseline showed lower attrition rates in comparison with those who did not meet these guidelines [23,24]. Davis and Addis [25] argued for investigation into the psychological determinants of behavior as predictors of attrition. Users with a low intention to change behavior have already been shown to drop out more often [26,27]. Accordingly, the role of other psychological determinants as predictors of attrition has not yet been fully explored.

Third, why do users stop using the intervention? Answers to this question may help researchers identify whether attrition is caused by features embedded in the intervention (eg, design of the intervention or technical problems with the intervention, lack of useful intervention content, too much questionnaires) or by reasons outside the intervention (eg, no interest in the topic, medical or emotional problems, lack of time).

In summary, the aim of this paper is threefold. The first aim is to examine when users stop using the intervention. The second aim is to investigate which users stop using the intervention informed by demographic variables, behavioral status at the beginning of the intervention, and psychological determinants. The third aim is to explore why users stop using the intervention by describing reasons for noncompletion.

This paper addresses these questions through secondary analysis of a digital health intervention that aimed to increase PA or reduce SB among the general population [28,29]. This intervention was developed using the health action process approach (HAPA) model, which describes psychological determinants that guide individuals in changing their behavior [30]. It is a 2-phase model that includes (1) motivational processes identified by determinants such as risk perception, outcome expectancies, and self-efficacy leading to a behavioral intention and (2) volitional processes identified by determinants such as action planning, coping planning, and self-monitoring bridging the gap between intention and the actual behavior [30]. As HAPA has been shown to effectively change behavior, the HAPA-based psychological determinants are considered important predictors of behavior change [30,31]. These predictors might not only influence behavior change but also the decision of whether to stop using an intervention.

Methods

Data Source

The data reported in this paper were from the MyPlan 2.0 factorial randomized controlled trial registered at ClinicalTrials.gov [NCT03274271] and approved by the Ghent University Hospital Ethics Committee. The protocol of the trial can be found elsewhere [28].

XSL•FO
RenderX

Intervention

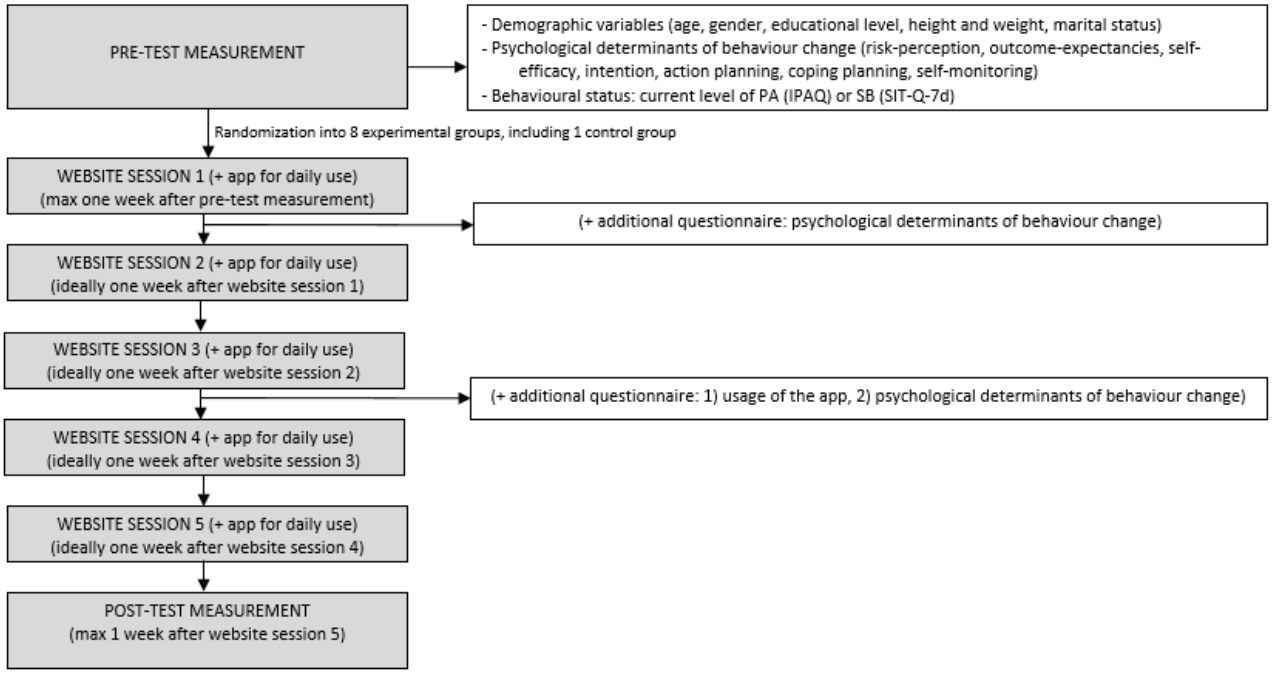
The MyPlan 2.0 digital health intervention consisted of a website and an optional mobile app to promote PA or reduce SB in healthy adults from the general population. MyPlan 2.0 was based on the HAPA model and consisted of a number of behavior change techniques (BCTs) aiming to influence participants' HAPA-based psychological determinants of behavior change. The BCTs used in this study were goal setting, providing information on consequences of behavior, providing feedback on performance, social support, action planning, coping planning, self-monitoring, and reviewing behavior goals. These BCTs are described below.

Before the start of the intervention, participants chose which behavior (PA or SB) they wanted to improve (ie, goal setting). Depending on their choice, they were directed to the version of MyPlan 2.0 targeting PA or SB. The structure of the intervention was identical for the two behaviors.

The website is considered the main part of the intervention and consisted of 5 website sessions, with 1 week between each session (see Figure 1). Participants were expected to go through each of these sessions. The structure of the website sessions was fixed. In the first session, participants created a profile, were offered an optional quiz with information about the benefits of the selected target behavior (ie, providing information on consequences of behavior), and received tailored feedback on the current state of their chosen behavior (ie, providing feedback on performance). Thereafter, participants created an action plan by specifying how they wanted to reach their PA or SB goal, what they wanted do to, and where and when they wanted to do it (ie, action planning). Consequently, they identified potential barriers and thought about possible solutions (ie, coping planning). Thereafter, participants were prompted to monitor their behavior via the app or other options such as writing in their diary or on their calendar (ie, self-monitoring). At the end of the first session, they could read about how they could obtain social support from their partner, friends, family, or colleagues. In the 4 follow-up sessions, participants were asked to reflect on their progress of behavior change of the past week by evaluating their PA or SB goal (ie, reviewing behavior goals). They were also prompted to adapt or maintain their action plan, coping plan, and self-monitoring method. Screenshots of the website can be found in Multimedia Appendix 1.

The app was offered to participants as an optional tool to provide support on a daily basis. The app was synchronized with the website and developed as an extension to support users with their plans created in the website sessions. Use of the app was not mandatory. It consisted of 5 modules through which participants could freely navigate. In the first module, participants could again obtain a quiz regarding the benefits of more PA or less SB. In the second module, participants could review their action plan (which was created on the website) and change their plan throughout the week (ie, action planning). Moreover, the app reminded participants of their plan by sending notifications at scheduled times. In the third module, they could select barriers and receive an overview of possible solutions (ie, coping planning). In the fourth module, participants received a notification every evening to monitor their behavior by rating if they succeeded in their plan for the day on a scale from 0 to 5 (ie, self-monitoring). In the fifth module, users could collect medals by completing the website sessions, completing quizzes, and monitoring their behavior. These elements of gamification (ie, "the use of game design elements in nongaming contexts" [32]) were added to increase engagement with the intervention [33]. Screenshots of the app can be found in Multimedia Appendix 2.

Figure 1. Flowchart of the MyPlan 2.0 intervention. IPAQ: International Physical Activity Questionnaire; PA: physical activity; SB: sedentary behavior; SIT-Q-7d: Last 7-day Sedentary Behavior Questionnaire.



XSL•FO
RenderX

Intervention Content as Part of the Design

Intervention content differed as part of the design of the MyPlan 2.0 factorial randomized controlled trial [28]. Participants were randomly allocated to 8 different groups to evaluate the efficacy of 3 BCTs (ie, action planning, coping planning, and self-monitoring) and their combinations. As such, each group received a different version of the intervention, in which the 3 different BCTs were combined (Table 1). In both the website and the app, the BCTs could easily be removed or added in order to create the different groups [28]. Nevertheless, each participant received a basic intervention including the following BCTs: goal setting, providing information on consequences of behavior, providing feedback on performance, social support, and reviewing behavior goals [29].

Table 1. Different intervention content for each group as part of the design of the MyPlan 2.0 factorial randomized controlled trial.

	Action planning	Coping planning	Self-monitoring
Group 1	+[a]	+	+
Group 2	+	+	–[b]
Group 3	+	–	+
Group 4	–	+	+
Group 5	+	–	–
Group 6	–	+	–
Group 7	–	–	+
Group 8	–	–	–

[a]+: group received the intervention content including the behavior change technique.

[b]–: group received the intervention content without the behavior change technique.

Participants and Procedure

The sample consisted of 473 participants who were recruited between February and December 2018 at the city library of Ghent or through social media. Inclusion criteria were a minimum age of 18 years, speaking Dutch, having internet access at home or work, and owning a smartphone (iOS or Android). Participants completed the 7 items of the Physical Activity Readiness Questionnaire as a screening instrument to detect individuals at risk for adverse effects when being more physically active [34]. Participants who answered no to all items were eligible for the study.

The flowchart for MyPlan 2.0 can be found in Figure 1. Participants completed pretest measurements including demographic variables, psychological determinants of behavior change, and questions assessing their current PA or SB level. When the pretest measurements were completed, participants were randomly allocated to 1 of the 8 different versions of the intervention (Table 1). Immediately after the randomization, participants could start with the intervention (maximum 1 week after the pretest measurements). The intervention consisted of 5 consecutive website sessions, ideally with 1 week between each session, and the optional mobile app, which could be used at any time during the intervention. Approximately 1 week after completing the last session, participants completed posttest measurements. The pretest measurements and posttest measurements were conducted via an online survey tool (Limesurvey GmbH).

Boosting strategies were used to encourage completion of each part (pretest and posttest measurements and the 5 website sessions): participants who did not complete a certain part after 1 week were sent a reminder, if they had not completed the part after 2 weeks, they were contacted by phone by the researcher (HS). If there was no response after 3 weeks, the participant was considered a noncompleter, and attrition was documented to have occurred during that specific part of the intervention. As such, the duration of the study could be different for each participant, depending on when they completed each part of the study.

After finishing website sessions 1 and 3, participants were invited to complete an additional questionnaire assessing psychological determinants during the intervention. After completing session 3, participants were also asked to complete another questionnaire assessing the use of the app [28]. Noncompletion of these additional questionnaires did not affect the continuation of the intervention. Data from these questionnaires were not used for analyses in this study.

Measures

Definition of Attrition

In this paper, we differentiated between 2 types of attrition: (1) nonusage attrition, which refers to participants who were not using the intervention (ie, not completing the website sessions) and (2) dropout attrition, which refers to participants who were lost to follow-up because they stopped completing questionnaires for research purposes (ie, did not complete posttest measurements). Here, nonusage attrition automatically equaled dropout attrition because of the linear design of the study (eg, it was not possible to start, for example, with session 4 on the website if session 3 was not completed). Consequently, in this paper, we will just use the term attrition. Not using the app was not considered attrition because participants could use the app as an optional choice. Moreover, not completing the additional questionnaires after website sessions 1 and 3 was not considered attrition because it was still possible to proceed with the online website sessions.

XSL•FO
RenderX

Demographic Variables

At the pretest measurement, the following demographic variables were assessed: age, gender, education level (categorized as not having vs having a college/university degree), BMI (categorized as not overweight [≤25 kg/m²] vs overweight [>25 kg/m²]), and marital status (categorized as not having a partner vs having a partner).

Behavioral Status

The current level of PA or SB was assessed at pretest measurement to determine the behavioral status at the beginning of the intervention. For PA, the Dutch long version of the International Physical Activity Questionnaire [35] was used to measure moderate-to-vigorous intensity PA (MVPA) in minutes per week. For SB, the Dutch 7-day sedentary behavior self-report questionnaire [36] was used to measure total sedentary time in hours per day. Behavioral status at the beginning of the intervention was categorized as not meeting the guidelines (<150 minutes per week of MVPA or >8 hours per day of sitting time) versus meeting the guidelines (>150 minutes per week of MVPA or <8 hours per day of sitting time).

Psychological Determinants

The HAPA-based psychological determinants were measured at pretest measurement using a set of 26 items (ie, at least 3 items per determinant), which can be found in Multimedia Appendix 3. As described in our protocol paper [28], the set was based on the HAPA model and was iteratively developed and validated by an expert panel using cognitive interviewing [37,38] and a discriminant content validity method [39]. The same set of items was used for the 2 behaviors but the items were adapted to either PA or SB. Risk perception was assessed by 4 items; 1 of the items was "I am a person who is prone to high blood pressure." The 4 items showed poor internal consistency (α=0.59). Removing the last item ("I am a person who is prone to have depression") increased the internal consistency (α=0.71). Therefore, only the first 3 items were used to assess risk perception. Outcome expectancies were assessed with 5 items; 1 of the items for PA was "If I start being physically active regularly, I will feel better afterward." Also here, the internal consistency was low (α=0.56). Removing the last item ("If I am physically active regularly, I have the feeling I lose time") increased the internal consistency (α=0.68). As a result, only the first 4 items were taken into account to assess outcome expectancies. Self-efficacy was assessed by 5 items; 1 of the items for SB was "I am sure I can reduce my sitting time, even when I feel tired." The items showed good internal consistency (α=0.83). Three items were used to assess intention; 1 of the items for PA was "I intend to be physically active regularly." The internal consistency for these items was good (α=0.87). Action planning was assessed by 3 items; 1 of the items for PA was "I know exactly what to do (how, where, when, ...) to be physically active regularly." All items showed good internal consistency (α=0.84). For coping planning, 3 items were used; 1 of the items for PA was "I already have thought about possible solutions in case I encounter obstacles in order to be physically active regularly (eg, if the swimming pool is closed, I go for a walk instead)." Also here, the items showed good internal consistency (α=0.88). Finally, 3 items were used to assess self-monitoring; 1 of the items for SB was "I am constantly monitoring how long I sit." The internal consistency for these items was good (α=0.76). Participants rated all items on a 5-point response scale (1=totally disagree, 2=somewhat disagree, 3=neutral, 4=somewhat agree, 5=totally agree). For each determinant, the mean score of the items was used in the analyses.

Attrition-Related Questionnaire

When a participant was determined to be a noncompleter of a certain intervention part, a questionnaire with reasons for discontinuation was sent by email. Participants could indicate whether they found the reason for attrition totally not applicable, not applicable, neutral, applicable or totally applicable in response to 8 statements concerning attrition. The questions were based on attrition-related factors described in an article by Eysenbach [10].

Statistical Analyses

Attrition Pattern

To analyze when users stopped using the intervention (aim 1), the numbers of participants per part of the intervention were described. For this paper, the different parts included the pretest and posttest measurements as well as the 5 website sessions (7 parts in total, see Figure 1). Each part was considered completed if participants completed the last question (for the pretest and posttest measurements) or visited the last page on the website (for the website sessions). Descriptive analyses were performed in Excel (Microsoft Corp).

Predictors of Attrition

To analyze which users stopped using the intervention, the following predictors of attrition were investigated: demographic variables, behavioral status, and psychological determinants at pretest measurement. Analyses were performed in SPSS (version 26, IBM Corp). Logistic regression models were fitted with attrition as a dependent variable at different time points (the number of the logistic regression models depended on the attrition pattern of aim 1). All independent variables (demographic variables, behavioral status, and psychological determinants) were entered separately into the logistic regression models. $P<.05$ was considered statistically significant, whereas P values between .05 and .10 were considered borderline significant; 95% confidence intervals were also reported.

In order to investigate whether different intervention content as part of the design of MyPlan 2.0 was a reason for attrition, 2 other predictors were added to the logistic regression models described above: the group to which the participants were allocated (group 1-8) and the choice of behavior participants wanted to improve (PA versus SB).

Reasons for Attrition

To analyze why users stopped using the intervention, the scores of participants to the attrition relation questionnaire were used. For each question, the number and percentage of participants who found the question (totally) not applicable, neutral, or (totally) applicable was shown. Descriptive analyses were performed in Excel (Microsoft Corp).

XSL•FO
RenderX

Results

Participant Characteristics

In total, 473 participants agreed to participate in the MyPlan 2.0 trial, completed pretest measurement, and were therefore considered users in this study. Of these participants, the mean age was 36.7 (SD 16.3) years, 69.1% (327/473) of participants were female, 66.4% (314/473) had a high level of education (college or university degree), 30.2% (143/473) were overweight, 42.3% (200/473) had a partner, and 49.0% (232/473) met the guidelines for either PA or SB. Descriptive statistics of the psychological determinants at pretest measurement are provided in Table 2. In addition, the number and percentage of participants in each group as well as the percentage of participants who chose PA or SB is provided (Table 2).

Table 2. Characteristics of participants of the factorial randomized controlled trial MyPlan 2.0.

Variable	Participants (n=473)
Age (years), mean (SD)	36.7 (16.3)
Gender (female), n (%)	327 (69.1)
Level of education (% high = university/college)	314 (66.4)
BMI (kg/m^2), mean (SD)	23.5 (3.7)
Overweight, n (%)	143 (30.2)
Marital status (with partner), n (%)	200 (42.3)
Behavioral status (meets guidelines of PA[a] or SB[b]), n (%)	232 (49.0)
Self-efficacy[c], mean (SD)	3.54 (0.62)
Outcome expectancies[c], mean (SD)	3.95 (0.49)
Risk perception[c], mean (SD)	2.08 (0.66)
Intention[c], mean (SD)	4.08 (0.55)
Action planning[c], mean (SD)	2.84 (0.81)
Coping planning[c], mean (SD)	2.47 (0.82)
Self-monitoring[c], mean (SD)	2.06 (0.85)
Participants in each group[d], n (%)	
Group 1	59 (12.5)
Group 2	60 (12.7)
Group 3	56 (11.8)
Group 4	56 (11.8)
Group 5	61 (12.9)
Group 6	59 (12.5)
Group 7	61 (12.9)
Group 8	61 (12.9)
Choice of behavior (participants who chose PA), n (%)	335 (70.8)

[a]PA: physical activity.

[b]SB: sedentary behavior.

[c]Mean on a score of 5 (SD).

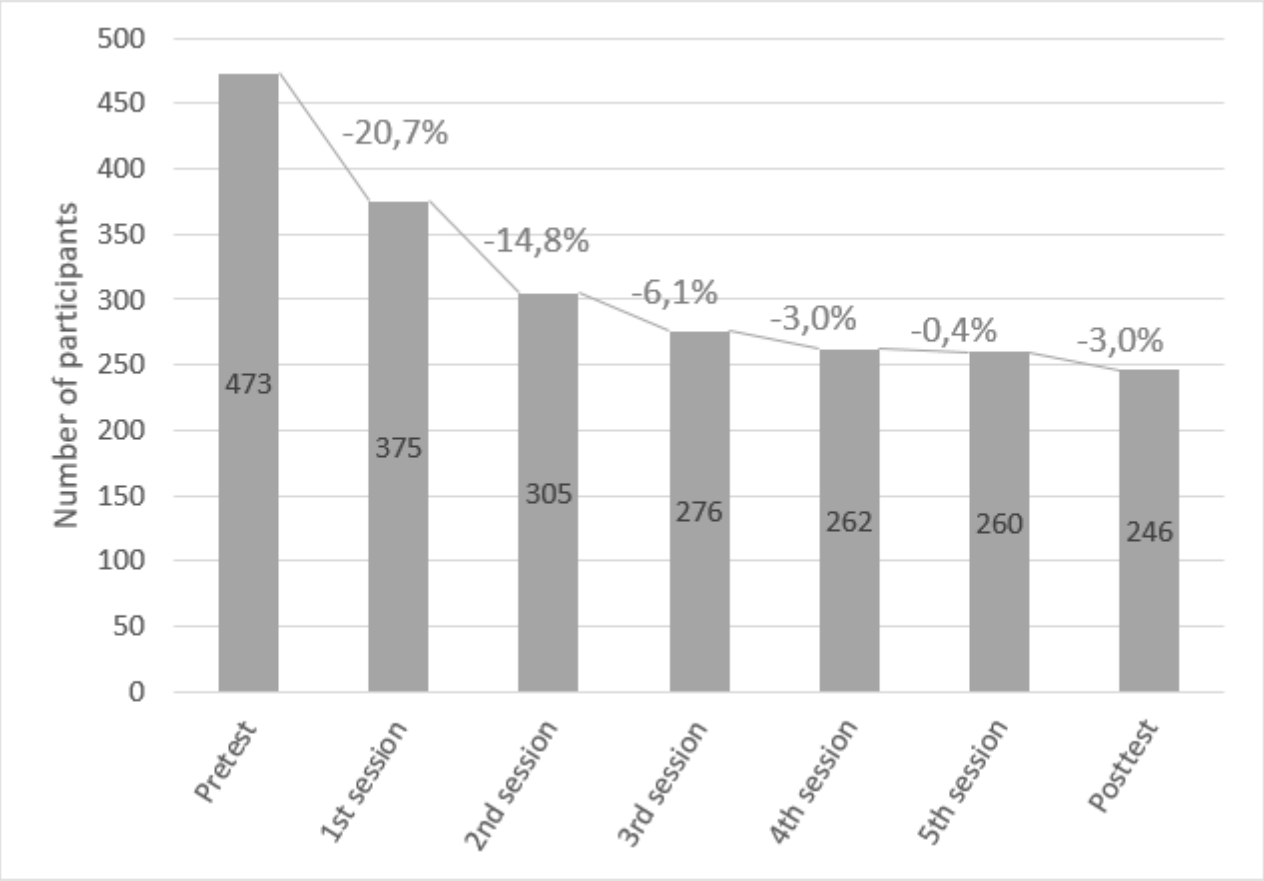
[d]See Table 1 for information about each group.

Attrition Pattern

Of the participants, 47.9% (227/473) did not complete the intervention. Figure 2 shows the attrition pattern of the intervention. The biggest loss of participants was found in the early stage of the intervention; 20.7% (98/473) dropped out before completing the first website session, 14.8% (70/473) before completing the second session, and 6.1% (29/473) before the third session. This means that 41.6% (197/473) of participants dropped out before the third session and only 6.3% (30/473) after that session, which could determine a steady state of attrition after that part of the intervention.

XSL•FO
RenderX

Figure 2. Attrition pattern of the intervention.



Predictors of Attrition

Predictors of attrition were investigated at 3 time points (Table 3). In order to do so, 3 logistic regression models were fitted: (1) identification of the predictors of first session completion, (2) identification of predictors of third session completion, and (3) identification of predictors of whole intervention completion (ie, completion of all 7 parts). This decision was based on the attrition pattern of aim 1: a large number of participants did not complete the first session, and a steady state of attrition was found after third session completion.

XSL•FO
RenderX

Table 3. Predictors of attrition.

Characteristics	First session completion (0=dropout before first session, 1=first session completion), OR[a] (95% CI)	*P* value	Third session completion (0=dropout before third session, 1=third session completion), OR (95% CI)	*P* value	Whole intervention completion (0=dropout before posttest measurements, 1=whole intervention completion), OR (95% CI)	*P* value
Demographic variables						
Age	1.0 (1.0-1.0)	0.91	1.0 (1.0-1.1)	0.21	1.0 (1.0-1.0)	0.16
Gender (0=female, 1=male)	1.4 (0.8-2.3)	0.20	1.4 (1.0-2.2)	0.08	1.4 (1.0-2.1)	0.07
Education (0=no college/university degree, 1=college/university degree)	1.2 (0.8-1.9)	0.46	1.2 (0.8-1.7)	0.46	1.0 (0.7-1.5)	0.89
BMI (0=not overweight, 1=overweight)	0.9 (0.5-1.4)	0.57	0.8 (0.6-1.2)	0.38	0.9 (0.6-1.3)	0.52
Marital status (0=no partner, 1=partner)	1.3 (0.8-2.1)	0.23	1.1 (0.8-1.6)	0.64	1.0 (0.7-1.5)	0.95
Baseline norm (0=did not meet guidelines, 1=met guidelines)	1.1 (0.7-1.7)	0.81	1.0 (0.7-1.4)	0.80	1.0 (0.7-1.5)	0.95
Psychological determinants						
Self-efficacy	1.0 (0.7-1.5)	0.86	1.2 (0.9-1.6)	0.30	1.1 (0.8-1.5)	0.54
Outcome expectancies	1.1 (0.7-1.7)	0.62	1.1 (0.8-1.5)	0.63	1.0 (0.7-1.4)	0.84
Risk perception	0.9 (0.7-1.2)	0.40	0.9 (0.7-1.2)	0.64	0.9 (0.7-1.2)	0.44
Intention	1.0 (0.7-1.5)	0.98	1.0 (0.7-1.4)	0.98	1.0 (0.7-1.4)	0.84
Action planning	0.8 (0.6-1.0)	0.08	0.8 (0.7-1.0)	0.09	0.9 (0.7-1.1)	0.25
Coping planning	0.8 (0.6-1.0)	0.05	0.9 (0.8-1.2)	0.64	1.0 (0.8-1.2)	0.65
Self-monitoring	0.8 (0.6-1.0)	0.19	0.8 (0.6-1.0)	0.04	0.8 (0.7-1.0)	0.08
Intervention content as part of the design						
Group	1.0 (0.9-1.1)	0.45	1.1 (1.0-1.1)	0.22	1.1 (1.0-1.1)	0.18
Behavior choice (0=participant chose PA[b], 1=participant chose SB[c])	1.2 (0.7-1.9)	0.52	1.3 (0.8-1.9)	0.26	1.2 (0.8-1.8)	0.29

[a]OR: odds ratio.

[b]PA: physical activity.

[c]SB: sedentary behavior.

Predictors of First Session Completion

No significant predictors of first session completion were found. However, the psychological determinants action planning and coping planning were found to be borderline significant (odds ratio [OR] 0.782 [95% CI 0.595-1.028], *P*=.08 and OR 0.769 [95% CI 0.590-1.002], *P*=.05, respectively), with participants with a higher score on action planning and coping planning being less likely to complete the first website session (Table 3).

Furthermore, the group to which participants were allocated and the choice of behavior participants wanted to improve as part of the design of MyPlan 2.0 were not significant predictors of first session completion.

Predictors of Third Session Completion

The psychological determinant self-monitoring significantly predicted whether participants completed the third session (OR 0.801 [95% CI 0.646-0.993], *P*=.04), with participants with a higher score on self-monitoring being less likely to complete the third website session. Furthermore, the demographic variable gender and the psychological determinant action planning were found to be borderline significant (OR 1.440 [95% CI 0.963-2.155], *P*=.08 and OR 0.822 [95% CI 0.655-1.032], *P*=.09, respectively). Men were more likely to complete the third website session, and participants with a higher score on action planning were less likely to complete the third website session (Table 3).

Furthermore, the group to which participants were allocated and the choice of behavior participants wanted to improve as part of the design of MyPlan 2.0 were not significant predictors of third session completion.

Predictors of Whole Intervention Completion

The demographic variable gender and the psychological determinant self-monitoring were found to be borderline significant (OR 1.437 [95% CI 0.969-2.13], *P*=.07 and OR 0.828 [95% CI 0.669-1.025], *P*=.08, respectively). Men were more likely to complete the whole intervention, and participants with a higher score on self-monitoring were less likely to complete the whole intervention (Table 3).

XSL•FO
RenderX

Furthermore, the group to which participants were allocated and the choice of behavior participants wanted to improve as part of the design of MyPlan 2.0 were not significant predictors of whole intervention completion.

Reasons for Attrition

The reasons why participants stopped using the intervention were obtained from 51 of 227 participants (22% of all noncompleters) and can be found in Table 4. We were not able to contact the other participants. Participants who were older (OR 1.021 [95% CI 1.002-1.041]) and had a higher educational level (OR 2.938 [95% CI 1.345-6.418]) were more likely to complete the questionnaire.

The most endorsed reasons to stop using the intervention were "Filling out the questionnaires took a lot of my time" (28/51, 55%), "I don't have time" (26/51, 50%), "The intervention doesn't provide useful content" (21/51, 41%), "I experienced technical problems with the website or app" (20/51, 39%), "I already meet the health guidelines for PA/SB" (16/51, 31%), and "I experienced medical/emotional problems" (8/51, 16%).

The following reasons were reported not to be important to stop using the intervention: "I am not interested in the topic" (1/51, 2%) and "I don't want to change my behavior" (1/51, 2%).

Table 4. Reasons for attrition (n=51).

I stopped using the interventions because...	(Totally) not applicable, n (%)	Neutral, n (%)	(Totally) applicable, n (%)
I am not interested in the topic	46 (90)	4 (8)	1 (2)
I don't have time	19 (37)	6 (12)	26 (51)
I already meet the health guidelines for PA[a]/SB[b]	19 (37)	16 (31)	16 (32)
I don't want to change my behavior	41 (80)	9 (18)	1 (2)
The intervention doesn't provide useful content	25 (49)	5 (10)	21 (41)
Filling out the questionnaires took a lot of my time	16 (31)	7 (14)	28 (55)
I experience technical problems with the website or app	28 (55)	3 (6)	20 (39)
I experience medical/emotional problems	42 (82)	1 (2)	8 (16)

[a]PA: physical activity.

[b]SB: sedentary behavior.

Discussion

Principal Findings

This study investigated when, which, and why users stopped using an intervention to promote PA and reduce SB. The study demonstrated that 227 of 473 participants stopped using the intervention, and drop out occurred mainly in the first weeks. Certain predictors of first session, third session, or whole intervention completion were found. The most endorsed reasons to stop using the intervention were time-consuming nature of the questionnaires, not having time, dissatisfaction with the content of the intervention, technical problems, already meeting the guidelines for PA/SB, and, to a lesser extent, experiencing medical/emotional problems.

Although our intervention was systematically developed based on both qualitative and quantitative research [40,41] and boosting strategies were used to keep participants engaged with the intervention, an overall attrition rate of 47.9% was observed. This rate is similar to other interventions [10,13]. Attrition is an important obstacle in digital health interventions and should be reduced in order to increase the public health impact. The overall pattern of results indicates that the largest group of users drop out in the first weeks of the intervention, and a stable group of users remains after that. This pattern is frequently observed in digital health research [10,42] and is described as a L-shaped curve [10]. A possible explanation for this attrition pattern might be the time-consuming nature of the questionnaires, and participants reported this as a reason to quit the intervention. Indeed, at the start of the intervention, participants completed several questionnaires, mainly for research purposes and for tailoring advice throughout the intervention. Although we reduced the amount of questions substantially in comparison with a previous version of the intervention (MyPlan 1.0) to prevent attrition [18,43], participants still perceived it as too long. The review by Sharpe et al [44] showed that users are indeed less inclined to persevere with digital health interventions when they are found to be time-consuming and burdensome. Researchers should thus thoroughly reflect which and how many questions should be included in digital health intervention studies. Another option might be to collect baseline data through monitoring devices (eg, wearables such as Fitbit). Although this is often done for research purposes [45,46], such devices can also be used to provide tailored support in the beginning of an intervention (eg, tailored feedback on their current PA level).

According to Eysenbach [10], attrition might also be the result of a wrong user group, the members of which quickly lost interest. Indeed, the overall pattern of results indicates that participants already doing action planning, coping planning, and self-monitoring were more likely to drop out. As the MyPlan 2.0 intervention focused on these postintentional determinants [28], the intervention may not have added value for these participants as they might have been the wrong user group, causing them to stop using the intervention. However, one should be reminded that some of these effects were borderline significant in the current analyses, and thus await further replication and corroboration. Notwithstanding, an important

XSL•FO
RenderX

question to answer is "Do participants already doing action planning, coping planning, and self-monitoring still benefit from an intervention?" On the one hand, one may reason that individuals who already have the competencies and skills to change behavior by themselves may not need additional support. On the other hand, it may well be that these individuals require a different, more individual approach that takes into account the needs and characteristics of the individual. Innovations in digital technology and artificial intelligence [47] enable researchers to develop more personalized interventions, making such an individual approach possible. Indeed, various studies indicate that personalization is crucial for future digital health interventions to increase engagement [44,48,49]. Yet, personalization can occur on different levels [48], and to the best of our knowledge, there is no consensus or framework on how to specifically personalize digital health interventions for PA or SB. Based on our findings, interventions may be personalized on 2 levels: dynamic tailoring of BCTs to the motivational stage to which an individual belongs and including personalized suggestions of BCTs at the operational level.

Regarding the first level, individuals may differ in terms of motivational stages: preintenders are individuals who do not yet have an intention to change, intenders are individuals who have an intention but do not yet act on these intentions, and actors are individuals who already act on their intentions [50]. Tailoring interventions to the stage of the participant may be more successful than mismatched interventions [51], but this tailoring often occurs only once at the beginning of an intervention. However, stages can also differ over time during the intervention (eg, intenders can become actors). By extension, research suggests [52,53] further differentiating between actors (individuals who recently started to perform the behavior) and maintainers (individuals who perform the behavior with high automatization over a long period of time). Intenders and actors may still benefit from BCTs such as action planning, coping planning, and self-monitoring, whereas maintainers might need other BCTs [29,52,54]. The findings of Schwarzer et al [55] indeed show that habitual activity does not require planning because the activity occurs rather automatically, whereas in the absence of the habit, planning appears to be a facilitator of PA. As such, providing dynamic tailoring of BCTs to these changing demands could reduce dropout in future interventions.

Determining the motivational stage of an individual is not an easy endeavor [56]. One might argue that individuals should be matched to stages based on meeting the health guidelines (eg, meeting the health guidelines may reflect being a maintainer and not meeting the health guidelines but having plans to work toward them may reflect being an intender). Accordingly, "Already meeting the guidelines for PA/SB" was one of the main reasons participants indicated stopping the intervention, with participants possibly needing other BCTs. However, meeting or not meeting the guidelines for PA/SB alone does not necessarily reflect the stage of the individual. One might also argue that individuals should be matched to stages based on the goal they have in mind. One can be a maintainer for a small behavior goal (eg, walking twice a week to work) and still be an intender or actor for a more challenging behavior goal (eg, running twice a week).

This brings us to the second level of personalization: including personalized suggestions of BCTs at the operational level. Participants who dropped out in our study might have been actors or maintainers who were looking for more challenging support. However, participants in our study were their own expert in terms of making action and coping plans, which means they had full control over the content of their plans. Although this is in line with self-regulation theory and increases autonomy [57], the delivery of these BCTs remained abstract and generic, offering standard but not personalized support. Indeed, it could be that participants were limiting themselves to plans that were already familiar to them, whereas they actually needed personalized suggestions that could provide them with new information and inspiration. Accordingly, dissatisfaction with intervention content was a reason for attrition in this study. This is in line with other research investigating user engagement [44]; participants in digital interventions for weight management most disliked generic information and repetition of content. As such, there is a need to tailor support at the operational level, involving suggestions of specific plans that are personalized to the individual. In addition, not having time was also found to be a reason for attrition. We acknowledge that thinking about action and coping plans is time intensive and requires high effort. Here, providing more personalized suggestions could result in a lower effort, time-effective intervention that could reduce dropout [58]. One should note, however, that behavior change in itself is not an easy endeavor, and raising awareness that behavior change takes time and effort is important. Here again, collecting objective data on PA/SB through monitoring devices [45] not only at the beginning but also throughout the course of the intervention will be important for personalization on both levels. That way, shifts between stages can be more easily identified (eg, intender to actor, actor to maintainer, actor to intender when there is a relapse). Passive data collection also offers the opportunity to provide more accurate personalized suggestions (eg, guiding participants from 8000 to 10,000 steps, guiding participants from walking to running, suggesting appropriate moments for a certain participant to do PA).

Remarkably, men were less likely to drop out than women. This is in contrast with previous findings [17,18,59]. A possible reason may be that this was an RCT compared to an open access study where participants had to give verbal consent for enrollment in the intervention. Several studies, including this one, have shown that men are less likely to enroll in studies compared to women [59], as women are more prone to respond in a socially desirable fashion [60]. However, once men do enroll in studies, they are more determined to complete the study. In order to increase engagement with the intervention, specific suggestions of plans as described in the previous paragraphs could also be personalized based on gender. Overall, most demographic variables did not predict whether certain subgroups of users stopped using the intervention. This could imply that the intervention can be broadly implemented and does not exclude specific target groups. Still, almost half of the participants dropped out of the study. This might indicate that other contextual and personal factors that were not investigated in this study (eg, social and physical environment, weather, location, mood, or health status) play a role.

XSL•FO
RenderX

Some other reasons not yet mentioned might explain why participants stopped using the intervention. First, although our website and app were thoroughly alpha and pilot tested [41], some technical problems were present at the beginning of the intervention that could have been a burden on participants. For example, our app did not work well with older smartphones, and the website caused technical problems when used in specific internet browsers (eg, Firefox did not always work well on tablets). Also, in the first weeks of the study, a bug caused the website to crash when a button for an optional website page was clicked on, preventing participants from returning to the main website page and completing their first website session. Future interventions should alpha test their websites and apps through all possible scenarios (various types of smartphones, internet browsers, laptops/tablets, etc) with a large user group. However, we should not assume that all technical problems can be solved in advance, as unforeseen barriers will always come up in digital health. Therefore, it may be useful if future interventions would provide a short manual with information to keep technical problems to a minimum (ie, press refresh when the website is stuck, use the internet browser Google Chrome, use the latest version of Android/iOS, do not use tablets). Second, some participants experienced medical/emotional problems during the intervention causing them to drop out. Future interventions should have the option to respond accordingly with particular advice or should refer to specific assistance (eg, doctor, psychologist, physiotherapist).

Strengths and Limitations

This study has several strengths. To the best of our knowledge, this is the first study that investigated multiple psychological determinants of behavior as predictors of attrition in an intervention to promote an active lifestyle. Many studies have already described predictors of attrition in digital health but focused mainly on demographic variables [17,19,20,22] or factors relating to the digital intervention itself instead of the health behavior (eg, attitudes toward the digital tool, perceived control over the tool) [61]. Second, this study investigated attrition in different parts of the intervention, whereas most studies only describe attrition rates at the end of their interventions [10]. Third, this study had a large study sample.

This study also has a number of limitations. First, only a small proportion of users reported reasons why they stopped using the intervention (22% of all noncompleters). This low response rate could be explained by the format used to investigate reasons for dropout (eg, although it was stated in our protocol paper that telephone calls would be used, an online questionnaire was used due to lack of time). Future studies still may consider the use of telephone calls. In addition, most of these users dropped out before the third session of the intervention, making it impossible to compare reasons for attrition at the beginning of the intervention with those at the end of the intervention. Second, considering the linear design of the study (see Methods), no posttest measurements of noncompleters were collected. As such, it was not possible to explore whether noncompleters improved their PA or SB levels due to their (short) participation in the intervention. Investigating this could be important in future studies, as stopping the intervention does not necessarily coincide with failure [62]. Third, as this is an RCT, this study could have shown different results if it would have been an open access study (where attrition rates are usually even higher) [24,63]. Fourth, the study sample consisted mostly of women (69.1%) and highly educated adults (66.4%), which has also been the case in other digital health intervention studies [64]. As such, one should be careful when generalizing the study outcomes to a broader population.

Conclusion

This study offered insights into when, which, and why users stop using a digital health intervention and provided some directions where future studies might focus on to prevent attrition. Personalization of interventions will be important, on one hand by dynamic tailoring of BCTs to the motivational stage to which an individual belongs and on the other hand by including personalized suggestions of BCTs at the operational level. Future studies should keep questionnaires (either for research purposes or tailoring) to a minimum by, for example, using objective monitoring devices, and technical aspects of digital health interventions should be thoroughly tested in advance.

Acknowledgments

We would like to thank Prof Dr Armand De Clercq, PhD, Louise Poppe, and Celien Van der Mispel for their support in developing MyPlan 2.0.

Authors' Contributions

All authors were involved in designing the study. HS collected and analyzed the data and drafted the manuscript. GC, IDB, and DVD critically revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the MyPlan 2.0 intervention website.
[PDF File (Adobe PDF File), 1030 KB - mhealth_v10i1e30583_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Screenshots of the MyPlan 2.0 mobile app.
[PDF File (Adobe PDF File), 526 KB - mhealth_v10i1e30583_app2.pdf]

Multimedia Appendix 3
A 26-item questionnaire on the Health Action Process Approach–based psychological determinants.
[PDF File (Adobe PDF File), 74 KB - mhealth_v10i1e30583_app3.pdf]

References

1. Müller A, Maher CA, Vandelanotte C, Hingle M, Middelweerd A, Lopez ML, et al. Physical activity, sedentary behavior, and diet-related ehealth and mHealth research: bibliometric analysis. J Med Internet Res 2018 Apr 18;20(4):e122 [FREE Full text] [doi: 10.2196/jmir.8954] [Medline: 29669703]
2. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
3. Glasgow RE, McKay H, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Educ Couns 2001 Aug;44(2):119-127. [doi: 10.1016/s0738-3991(00)00186-5] [Medline: 11479052]
4. Broekhuizen K, Simmons D, Devlieger R, van Assche A, Jans G, Galjaard S, et al. Cost-effectiveness of healthy eating and/or physical activity promotion in pregnant women at increased risk of gestational diabetes mellitus: economic evaluation alongside the DALI study, a European multicenter randomized controlled trial. Int J Behav Nutr Phys Act 2018 Mar 14;15(1):23 [FREE Full text] [doi: 10.1186/s12966-018-0643-y] [Medline: 29540227]
5. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, present, and future of eHealth and mHealth research to improve physical activity and dietary behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
6. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013 Jan;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
7. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-based physical activity interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
8. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
9. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
10. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
11. Poppe L, De Bourdeaudhuij I, Verloigne M, Shadid S, Van Cauwenberg J, Compernolle S, et al. Efficacy of a self-regulation-based electronic and mobile health intervention targeting an active lifestyle in adults having type 2 diabetes and in adults aged 50 years or older: two randomized controlled trials. J Med Internet Res 2019 Aug 02;21(8):e13363 [FREE Full text] [doi: 10.2196/13363] [Medline: 31376274]
12. Degroote L, Plaete J, De Bourdeaudhuij I, Verloigne M, Van Stappen V, De Meester A, et al. The effect of the eHealth intervention 'MyPlan 1.0' on physical activity in adults who visit general practice: a quasi-experimental trial. Int J Environ Res Public Health 2018 Jan 30;15(2):228 [FREE Full text] [doi: 10.3390/ijerph15020228] [Medline: 29385770]
13. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
14. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 Steps Australia. J Med Internet Res 2015;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]
15. Hutchesson MJ, Gough C, Müller A, Short CE, Whatnall MC, Ahmed M, et al. eHealth interventions targeting nutrition, physical activity, sedentary behavior, or obesity in adults: a scoping review of systematic reviews. Obes Rev 2021 Oct 23;22(10):e13295. [doi: 10.1111/obr.13295] [Medline: 34159684]
16. Tsay CH, Kofinas AK, Trivedi SK, Yang Y. Overcoming the novelty effect in online gamified learning systems: an empirical evaluation of student engagement and performance. J Comput Assist Learn 2019 Dec 04;36(2):128-146. [doi: 10.1111/jcal.12385]

XSL•FO
RenderX

17. Reinwand DA, Schulz DN, Crutzen R, Kremers SP, de Vries H. Who follows eHealth interventions as recommended? a study of participants' personal characteristics from the experimental arm of a randomized controlled trial. J Med Internet Res 2015;17(5):e115 [FREE Full text] [doi: 10.2196/jmir.3932] [Medline: 25963607]
18. Van der Mispel C, Poppe L, Crombez G, Verloigne M, De Bourdeaudhuij I. A self-regulation-based eHealth intervention to promote a healthy lifestyle: investigating user and website characteristics related to attrition. J Med Internet Res 2017 Jul 11;19(7):e241 [FREE Full text] [doi: 10.2196/jmir.7277] [Medline: 28698168]
19. Kelders SM, Van Gemert-Pijnen JEWC, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011 Apr;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
20. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020 Feb 17;3(1):21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
21. Wurst R, Maliezefski A, Ramsenthaler C, Brame J, Fuchs R. Effects of incentives on adherence to a web-based intervention promoting physical activity: naturalistic study. J Med Internet Res 2020 Jul 30;22(7):e18338 [FREE Full text] [doi: 10.2196/18338] [Medline: 32729835]
22. Reinwand DA, Crutzen R, Elfeddali I, Schneider F, Schulz DN, Smit E, et al. Impact of educational level on study attrition and evaluation of web-based computer-tailored interventions: results from seven randomized controlled trials. J Med Internet Res 2015 Oct 07;17(10):e228 [FREE Full text] [doi: 10.2196/jmir.4941] [Medline: 26446779]
23. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007 Jan 22;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
24. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
25. Davis MJ, Addis ME. Predictors of attrition from behavioral medicine treatments. Ann Behav Med 1999 Dec;21(4):339-349. [doi: 10.1007/BF02895967] [Medline: 10721442]
26. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013 Aug 30;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
27. Peels DA, Bolman C, Golsteijn RHJ, De Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between Web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res 2012 Dec 17;14(6):e179 [FREE Full text] [doi: 10.2196/jmir.2229] [Medline: 23246790]
28. Schroé H, Van der Mispel C, De Bourdeaudhuij I, Verloigne M, Poppe L, Crombez G. A factorial randomised controlled trial to identify efficacious self-regulation techniques in an e- and m-health intervention to target an active lifestyle: study protocol. Trials 2019 Jun 10;20(1):340 [FREE Full text] [doi: 10.1186/s13063-019-3456-7] [Medline: 31182147]
29. Schroé H, Van Dyck D, De Paepe A, Poppe L, Loh WW, Verloigne M, et al. Which behaviour change techniques are effective to promote physical activity and reduce sedentary behaviour in adults: a factorial randomized trial of an e- and m-health intervention. Int J Behav Nutr Phys Act 2020 Oct 07;17(1):127 [FREE Full text] [doi: 10.1186/s12966-020-01001-x] [Medline: 33028335]
30. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
31. Zhang C, Zhang R, Schwarzer R, Hagger MS. A meta-analysis of the health action process approach. Health Psychol 2019 Jul;38(7):623-637. [doi: 10.1037/hea0000728] [Medline: 30973747]
32. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. In: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. New York, NY: Association for Computing Machinery; 2011 Presented at: MindTrek '11: Academic MindTrek 2011; September 28-30; Tampere, Finland p. 9-15 URL: https://uwaterloo.ca/scholar/sites/ca.scholar/files/lnacke/files/From_game_design_elements_to_gamefulness-_defining_gamification.pdf [doi: 10.1145/2181037.2181040]
33. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One 2017 Mar;12(3):e0173403 [FREE Full text] [doi: 10.1371/journal.pone.0173403] [Medline: 28362821]
34. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
35. Vandelanotte C, De Bourdeaudhuij I, Philippaerts R, Sjöström M, Sallis J. Reliability and validity of a computerized and Dutch version of the International Physical Activity Questionnaire (IPAQ). J Phys Activity Health 2005 Jan;2(1):63-75. [doi: 10.1123/jpah.2.1.63]

XSL•FO
RenderX

36. Wijndaele K, DE Bourdeaudhuij I, Godino JG, Lynch BM, Griffin SJ, Westgate K, et al. Reliability and validity of a domain-specific last 7-d sedentary time questionnaire. Med Sci Sports Exerc 2014 Jun;46(6):1248-1260. [doi: 10.1249/MSS.0000000000000214] [Medline: 24492633]
37. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. Pub Opin Q 2007 Jun 05;71(2):287-311. [doi: 10.1093/poq/nfm006]
38. Christodoulou C, Junghaenel DU, DeWalt DA, Rothrock N, Stone AA. Cognitive interviewing in the evaluation of fatigue items: results from the patient-reported outcomes measurement information system (PROMIS). Qual Life Res 2008 Dec 12;17(10):1239-1246 [FREE Full text] [doi: 10.1007/s11136-008-9402-x] [Medline: 18850327]
39. Johnston M, Dixon D, Hart J, Glidewell L, Schröder C, Pollard B. Discriminant content validity: a quantitative methodology for assessing content of theory-based measures, with illustrative applications. Br J Health Psychol 2014 May 15;19(2):240-257. [doi: 10.1111/bjhp.12095] [Medline: 24628841]
40. Poppe L, Crombez G, De Bourdeaudhuij I, Van der Mispel C, Shadid S, Verloigne M. Experiences and opinions of adults with type 2 diabetes regarding a self-regulation-based eHealth intervention targeting physical activity and sedentary behaviour. Int J Environ Res Public Health 2018 May 10;15(5):954 [FREE Full text] [doi: 10.3390/ijerph15050954] [Medline: 29748460]
41. Poppe L, Van der Mispel C, Crombez G, De Bourdeaudhuij I, Schroé H, Verloigne M. How users experience and use an eHealth intervention based on self-regulation: mixed-methods study. J Med Internet Res 2018 Oct 01;20(10):e10412 [FREE Full text] [doi: 10.2196/10412] [Medline: 30274961]
42. Wangberg SC, Bergmo TS, Johnsen JK. Adherence in internet-based interventions. Patient Prefer Adherence 2008 Feb 02;2:57-65 [FREE Full text] [Medline: 19920945]
43. Poppe L, Van der Mispel C, De Bourdeaudhuij I, Verloigne M, Shadid S, Crombez G. Users' thoughts and opinions about a self-regulation-based eHealth intervention targeting physical activity and the intake of fruit and vegetables: a qualitative study. PLoS One 2017 Dec 21;12(12):e0190020 [FREE Full text] [doi: 10.1371/journal.pone.0190020] [Medline: 29267396]
44. Sharpe EE, Karasouli E, Meyer C. Examining factors of engagement with digital interventions for weight management: rapid review. JMIR Res Protoc 2017 Oct 23;6(10):e205 [FREE Full text] [doi: 10.2196/resprot.6059] [Medline: 29061557]
45. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110. [doi: 10.2196/jmir.9157] [Medline: 29567635]
46. Degroote L, De Bourdeaudhuij I, Verloigne M, Poppe L, Crombez G. The accuracy of smart devices for measuring physical activity in daily life: validation study. JMIR Mhealth Uhealth 2018 Dec 13;6(12):e10972 [FREE Full text] [doi: 10.2196/10972] [Medline: 30545810]
47. Mariano B. Towards a global strategy on digital health. Bull World Health Organ 2020 Apr 01;98(4):231. [doi: 10.2471/blt.20.253955]
48. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec 09;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
49. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 Mar 23:1 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
50. Lippke S, Ziegelmann JP, Schwarzer R. Stage-specific adoption and maintenance of physical activity: testing a three-stage model. Psychol Sport Exerc 2005 Sep;6(5):585-603. [doi: 10.1016/j.psychsport.2004.11.002]
51. Lippke S, Schwarzer R, Ziegelmann JP, Scholz U, Schüz B. Testing stage-specific effects of a stage-matched intervention: a randomized controlled trial targeting physical exercise and its predictors. Health Educ Behav 2010 Aug;37(4):533-546. [doi: 10.1177/1090198109359386] [Medline: 20547760]
52. Lippke S, Ziegelmann JP. Understanding and modeling health behavior: the multi-stage model of health behavior change. J Health Psychol 2006 Jan 01;11(1):37-50. [doi: 10.1177/1359105306058845] [Medline: 16314379]
53. Wienert J, Kuhlmann T, Storm V, Reinwand D, Lippke S. Latent user groups of an eHealth physical activity behaviour change intervention for people interested in reducing their cardiovascular risk. Res Sports Med 2019 Jul 26;27(1):34-49. [doi: 10.1080/15438627.2018.1502181] [Medline: 30047785]
54. Maher J, Conroy D. Habit strength moderates the effects of daily action planning prompts on physical activity but not sedentary behavior. J Sport Exerc Psychol 2015 Feb;37(1):97-107. [doi: 10.1123/jsep.2014-0258] [Medline: 25730895]
55. Schwarzer R, Warner L, Fleig L, Gholami M, Salvatore S, Cianferotti L, et al. Psychological mechanisms in a digital intervention to improve physical activity: a multicentre randomized controlled trial. Br J Health Psychol 2018 May 19;23(2):296-310. [doi: 10.1111/bjhp.12288] [Medline: 29265563]
56. Littell JH, Girvin H. Stages of change. A critique. Behav Modif 2002 Apr 26;26(2):223-273. [doi: 10.1177/0145445502026002006] [Medline: 11961914]
57. Maes S, Karoly P. Self-regulation assessment and intervention in physical health and illness: a review. Appl Psychol 2005 Apr;54(2):267-299. [doi: 10.1111/j.1464-0597.2005.00210.x]

XSL•FO
RenderX

58. Baumel A, Muench FJ. Effort-optimized intervention model: framework for building and analyzing digital interventions that require minimal effort for health-related gains. J Med Internet Res 2021 Mar 12;23(3):e24905 [FREE Full text] [doi: 10.2196/24905] [Medline: 33709943]
59. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007 May;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
60. Dalton D, Ortegren M. Gender differences in ethics research: the importance of controlling for the social desirability response bias. J Bus Ethics 2011 Apr 11;103(1):73-93. [doi: 10.1007/s10551-011-0843-8]
61. Wojtowicz M, Day V, McGrath PJ. Predictors of participant retention in a guided online self-help program for university students: prospective cohort study. J Med Internet Res 2013;15(5):e96 [FREE Full text] [doi: 10.2196/jmir.2323] [Medline: 23697614]
62. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
63. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
64. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595. [doi: 10.1093/her/cyp063] [Medline: 19897515]

Abbreviations

BCT: behavior change technique
HAPA: health action process approach
MVPA: moderate-to-vigorous physical activity
OR: odds ratio
PA: physical activity
SB: sedentary behavior

Edited by L Buis; submitted 20.05.21; peer-reviewed by A Mustafa, B Chaudhry, H Mehdizadeh; comments to author 30.07.21; revised version received 01.09.21; accepted 20.12.21; published 31.01.22.

Please cite as:
Schroé H, Crombez G, De Bourdeaudhuij I, Van Dyck D
Investigating When, Which, and Why Users Stop Using a Digital Health Intervention to Promote an Active Lifestyle: Secondary Analysis With A Focus on Health Action Process Approach–Based Psychological Determinants
JMIR Mhealth Uhealth 2022;10(1):e30583
URL: https://mhealth.jmir.org/2022/1/e30583
doi:10.2196/30583
PMID:

©Helene Schroé, Geert Crombez, Ilse De Bourdeaudhuij, Delfien Van Dyck. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 31.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Control of Personal mHealth Data Using a Mobile Blockchain App: Design Science Perspective

Arijit Sengupta[1*], BTECH, MS, PhD; Hemang Subramanian[1*], BTECH, MS, EMBA, PhD

Department of Information Systems and Business Analytics, College of Business, Florida International University, Miami, FL, United States

[*]all authors contributed equally

Corresponding Author:
Arijit Sengupta, BTECH, MS, PhD
Department of Information Systems and Business Analytics
College of Business
Florida International University
11200 Southwest 8th Street
Miami, FL, 33199
United States
Phone: 1 3053481427
Email: arijit.sengupta@fiu.edu

Abstract

Background: Integrating pervasive computing with blockchain's ability to store privacy-protected mobile health (mHealth) data while providing Health Insurance Portability and Accountability Act (HIPAA) compliance is a challenge. Patients use a multitude of devices, apps, and services to collect and store mHealth data. We present the design of an internet of things (IoT)–based configurable blockchain with different mHealth apps on iOS and Android, which collect the same user's data. We discuss the advantages of using such a blockchain architecture and demonstrate 2 things: the ease with which users can retain full control of their pervasive mHealth data and the ease with which HIPAA compliance can be accomplished by providers who choose to access user data.

Objective: The purpose of this paper is to design, evaluate, and test IoT-based mHealth data using wearable devices and an efficient, configurable blockchain, which has been designed and implemented from the first principles to store such data. The purpose of this paper is also to demonstrate the privacy-preserving and HIPAA-compliant nature of pervasive computing-based personalized health care systems that provide users with total control of their own data.

Methods: This paper followed the methodical design science approach adapted in information systems, wherein we evaluated prior designs, proposed enhancements with a blockchain design pattern published by the same authors, and used the design to support IoT transactions. We prototyped both the blockchain and IoT-based mHealth apps in different devices and tested all use cases that formed the design goals for such a system. Specifically, we validated the design goals for our system using the HIPAA checklist for businesses and proved the compliance of our architecture for mHealth data on pervasive computing devices.

Results: Blockchain-based personalized health care systems provide several advantages over traditional systems. They provide and support extreme privacy protection, provide the ability to share personalized data and delete data upon request, and support the ability to analyze such data.

Conclusions: We conclude that blockchains, specifically the consensus, hasher, storer, miner architecture presented in this paper, with configurable modules and software as a service model, provide many advantages for patients using pervasive devices that store mHealth data on the blockchain. Among them is the ability to store, retrieve, and modify ones generated health care data with a single private key across devices. These data are transparent, stored perennially, and provide patients with privacy and pseudoanonymity, in addition to very strong encryption for data access. Firms and device manufacturers would benefit from such an approach wherein they relinquish user data control while giving users the ability to select and offer their own mHealth data on data marketplaces. We show that such an architecture complies with the stringent requirements of HIPAA for patient data access.

(JMIR Mhealth Uhealth 2022;10(1):e32104) doi:10.2196/32104

KEYWORDS

blockchain; mobile apps; mining; HIPAA; personal health data; data privacy preservation; security; accuracy; transaction safety

XSL•FO
RenderX

Introduction

Background

Data is the new oil. [Clive Humby]

This quote by Clive Humby epitomizes the reality of today's internet-connected society, wherein private firms collect, distribute, store, analyze, and monetize user data. An important but understudied challenge facing the health care industry, users, and service providers of health care apps per se is the dichotomy of standards pertaining to user health care data. On the one hand, there are stringent data requirements such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR), which assure patient confidentiality, privacy, and security of user data, with most of the burden of data privacy and data security transferred to the provider. On the other hand, there are numerous pervasive devices collecting mobile health (mHealth) data such as height, weight, heart rate, electrocardiograms, sleep patterns, oxygen saturation levels, eye scans, blood pressure parameters, and real-time blood glucose levels with such data stored on devices and later transmitted to private clouds of equipment manufacturers and service providers at scale. Very little oversight or legal protection is provided to users of such devices and mHealth apps to prevent third parties from monetizing such personal data. An illustrious example is when Google recently released a machine learning algorithm that can actively predict a patient's heart condition by applying deep learning to retinal photographs [1]. In addition, user data are often harvested for research and analysis purposes, which often claim the anonymity of user data.

Although on the periphery, given a choice, >90% of the users surveyed in a recent study chose to keep their data private, and a significant number of them trusted to keep their health data on a blockchain [2]. A challenge facing health care providers, governments, and law enforcement is the cost of enforcing stringent requirements as per the HIPAA [3], GDPR, and the Public Health Emergency Privacy Act. For example, although HIPAA provides requirements for data privacy, the enforcement, monitoring, and penalizing of violators are practical difficulties [4]. The premise of supporting data privacy and access to personal health care data is limited when users sign onto health care apps under *terms and conditions* that prevent unauthorized access without the users' ability to *share* the data with their provider.

In this paper, we present a novel blockchain-centric approach to pervasive mHealth data by designing and deploying mobile apps that transmit and store data on a configurable blockchain optimized for storing, retrieving, and accessing internet of things (IoT) data. We summarize the blockchain architecture and how it supports IoT transactions using web services. We then develop 2 separate mHealth apps, one on iOS and the other on Android, and demonstrate the novelty of the blockchain-centric pervasive health care apps. We show how users of mHealth apps built atop the blockchain control their data while the blockchain supports access control, privacy, anonymity, and decentralized storage (not in the control of a single firm) [5].

Objective

To satisfy the challenges of user privacy and data access, combined with the need to maintain high security, speed, and availability of pervasive mHealth data, we propose using a configurable blockchain architecture such as the consensus, hasher, storer, miner (CHASM) architecture [6]. In this paper, we investigate the following research questions:

1. Research question 1: What and how do pervasive mobile apps interface with a configurable blockchain to provide users privacy, security, and control over their mHealth data?
2. Research question 2: How and can pervasive mobile apps adhere to stringent HIPAA compliance?

To our knowledge, this is one of the earliest researches to present a configurable blockchain architecture that combines elements of private and public blockchains and is compatible with the requirements of pervasive mHealth data, which caters to the highest plausible security and privacy requirements as prescribed by HIPAA. Access control is provided at multiple layers, that is, at the wallet layer, at the app layer, and at the blockchain layer. Similarly, prior research has not addressed how such pervasive health care data, which are generated by devices, can be made compatible with IoT blockchain data stores with significant security, high throughput, and expectations of low-synchronization times.

In the following sections, we describe the Design Science Research Methodology (DSRM) and then follow each step in the DSRM to implement a viable solution.

Methods

Overview

We used the DSRM, which is commonly used by researchers in information systems and computer science [7]. The main steps in the DSRM are listed in Textbox 1.

Following these principles, we executed the following steps for our solution: first, we defined the research problem and justified both our solution and the value of our solution; second, we defined the design goals for our system as the objectives that we intended to accomplish; third, we designed and implemented our solutions; fourth, we demonstrated the solution with respect to the goals set forth earlier; and finally, we evaluated the solution for accuracy, security, and (potential) costs while documenting the risks.

XSL•FO
RenderX

Textbox 1. Different steps of the Design Science Research Methodology and our approach to the solution.

- Phase 1
  - Step 1: problem definition and importance of solving
  - Step 2: identifying objectives of the solution
- Phase 2
  - Step 3: design and development of prototype
  - Step 4: demonstration of artifact
- Phase 3
  - Step 5: evaluation of artifact against requirements
  - Step 6: conclusion and communication

Step 1 of the DSRM: Problem Definition and Importance of Solving

Pervasive devices and apps on such devices that capture individual mHealth data have become popular in everyday use for millions of individuals using wearables, personal health parameter test kits, and medical devices [8]. Such devices and apps have several benefits by positively affecting patient health outcomes through the gamification of health care practices such as increasing the frequency of exercise, monitoring food intake and obesity control, improving communication among patients, and increasing patient motivation by encouraging them to join peer groups of similar individuals [9].

mHealth users generate a significant amount of data through their pervasive devices, which record data such as heart rate, blood cholesterol, and blood pressure. Most often, users own a multitude of devices such as smartwatches, smart apps on phones, smart blood sugar monitors, and smart brainwave readers [2,8,10-12]. Such pervasive devices create large data footprints, and such generated data are usually stored on separate and independent cloud-based servers or network-attached storage. Such databases are centrally administered, and data access is controlled by the firm that manages these data stores after the user's approval. A key issue with such centrally managed data stores is that users have no control over their data and frequently do not have access to historical data. However, very often, such data can provide valuable insights into user health, and when services such as those demonstrated by Google [1] or by Sleep City [13,14] become more commonplace, it can lead to early diagnosis of medical conditions or provide early warnings about the onset of diseases. Pervasive devices responsible for collecting mHealth data are called edge devices as they have exceedingly small storage capacity and depend on the network to collect and store data [15]. As a result of their resource-lessness and the need to store back-end data elsewhere, blockchains have been shown to provide numerous benefits [16]. However, such benefits are not automatically transferred over pervasive devices and apps that need to be specifically written with security, access, privacy, and performance considerations. Prior health care research on data on health care exchanges, data tamperproofing, and securing health care data have demonstrated the benefits of using blockchains in the context of health care [8].

Several public blockchains are used for building decentralized applications such as those in supply chain management, decentralized finance, and enabling contracts. The blockchain's advanced features of enabling private-public key cryptography, decentralized consensus algorithms that can be modified, and strong 1-way encryption of data through hashing can address several challenges facing health care data storage [16]. Fang et al [17] discussed the key challenges with current blockchain designs of public, consortium-based, and private blockchains in ensuring that all of the properties required for health care systems (ie, privacy, security, scalability, and immutability of patient health records) cannot be simultaneously available on the same blockchain. This is also known as blockchain trilemma.

The underlying blockchain we developed supports the symmetric encryption of transaction payloads in addition to signing. For this app, we used advanced encryption standards using a *secret key* associated with the user's private key. A user may associate multiple secret keys for different apps using the same private key. The mHealth data can only be decrypted using the user's secret key for onward sharing, downloading, and reading. A user can create multiple secret keys for each app, and each app can run on multiple devices to capture information about the same individual. For example, *secret key 1* associated with the user's ID would be used to capture the user's heart rate; *secret key 2* associated with the user's exercise log would be used to capture data on the number of steps a user takes while running or doing exercise; *secret key 3* could be used capture, for example, sleep patterns; and *secret key 4* could be used to log other health care information such as blood pressure, oxygen saturation levels in the blood, and blood sugar. Unique information would be stored in separate wallets or devices and could be deposited or synced up with the blockchain for overall persistent storage.

A feature of HIPAA compliance is providing users with absolute privacy of information and the ability to delete and clean out their data if given a chance. Textbox 2 presents the main requirements of such a software system and why each of these requirements is important to solve.

XSL•FO
RenderX

Textbox 2. A summary of the key requirements and why these are important to solve.

- Requirement 1: support for continuous data storage outside devices
  - Edge devices do not have sufficient storage. As a result, a network-based storage mechanism is needed.
- Requirement 2: diverse types of devices, data, and frequencies that transmit data
  - A person's mobile health care footprint is stored across multiple devices.
- Requirement 3: control access for distinct types of data
  - Each device generates a different kind of data and needs separate storage and access control if it were to be accessed by the user.
- Requirement 4: Health Insurance Portability and Accountability Act compliance
  - These standards governing health care data are essential to ensure that mobile health data are securely stored and that privacy is ensured.
- Requirement 5: stakeholder incentives for maintaining nodes, mining algorithm, and functionality of the blockchain
  - Stakeholders of the system can maintain and control the data on the blockchain. A portion of the revenues from device manufacturers and app sales would be used to fund maintenance, mining, and network maintenance.

Step 2 of the DSRM: Objectives of the Solution

A blockchain is a decentralized and distributed data store that can address the potential privacy and security concerns related to data access and data standardization while supporting device interoperability and providing users with total control of their own data [10,11]. Blockchain is a secure and immutable transaction ledger distributed on computing devices. Transactions are stored on the ledger through cryptographic validation and linked upward through the network. Owing to decentralization, peers can transact without a third party. Peers access the blockchain through a combination of private–public key cryptography (usually Elliptic Curve Digital Signature Algorithm) and can create and store as much data as they want within the blockchain transaction record. Such transactions are cryptographically signed by the owner of the data using his or her private key and are accessible on the blockchain only by the owner of the public–private key. All stakeholders of data, for example, device manufacturers, device users, marketplace administrators, and health care solution providers, can be permissioned onto the blockchain network [10], or users can explicitly or implicitly share data with other users.

The key advantage of using a blockchain system to track data is that the blockchain can verify and perennially store all created and validated user data. Some of the common advantages of using a blockchain versus a standard pervasive database system are listed in Table 1.

XSL•FO
RenderX

Table 1. Advantages of using blockchain for storing mobile health data.

Property of blockchain-based solution	Objectives accomplished for the mHealth app	Standard database or local storage–based apps
Anonymity	A user need not register with his or her personal identifiers and is associated with the data using the private key and public key information.	Access needs to be given by the administrator of the database.
Decentralization	The data is stored on a public infrastructure supported by individual users on a globally distributed network. In our prototype, we tested this with 4 parallel nodes.	Data is centralized and may be controlled by ≥1 administrators.
Transactional safety	Each transaction comprising data is signed with the user's private key, preventing others from manipulating the transaction.	Administrators are responsible for transactional safety. The database network subsystems are controlled by manufacturers.
Consistency	Irrespective of the type of data being sent on the network, the data is stored on the network as is without modification. In addition, the user may choose to encrypt the data before adding to the transaction payload and also for additional privacy.	The database can modify, alter, and change data when replicated. For example, in distributed databases, the design of the database could ensure that the data is compressed (potentially with loss of information).
Incentivization	Blockchain-based data marketplaces can enable users to be rewarded based on the validation of high-quality data that users can sell to other users.	Databases, by design, do not incentivize anyone. There is a central authority that decides all modes of access.
Perennial storage of data	Public blockchains provide a public space on the distributed ledger to store all kinds of data. With innovative architectures, it is possible for the storage of the data to also exist perennially on the blockchain.	Data can be deleted by administrators or anyone with administrative access.
Privacy preservation	Users who possess their own private key can access their data on the blockchain.	There is no privacy preservation by design. Administrators have all access rights and can give additional access to users.
Pervasive data access across multiple devices, apps, and systems	Users can access an infinite (theoretically) number of wallets, each of which can store a different type of health data. This provides a single-window clearance to all the user's health data.	Users are limited by the number of access accounts they are provided with by administrators in the system. User accounts are not anonymized either.
Ability to control access to data	The user who has private and public keys can control the data entirely. He or she is the only person that can create the data and access the data (if it is encrypted with his or her key).	Any administrator can control the data.
Ability to prevent access to user data (delete)	Within the blockchain and the pervasive apps, multiple approaches enable individuals to prevent access to the data available on the blockchain. For example, using multi-sig wallets enables data storage to not be accessible after the deletion of one key.	The access is controlled by the database administrator per se.

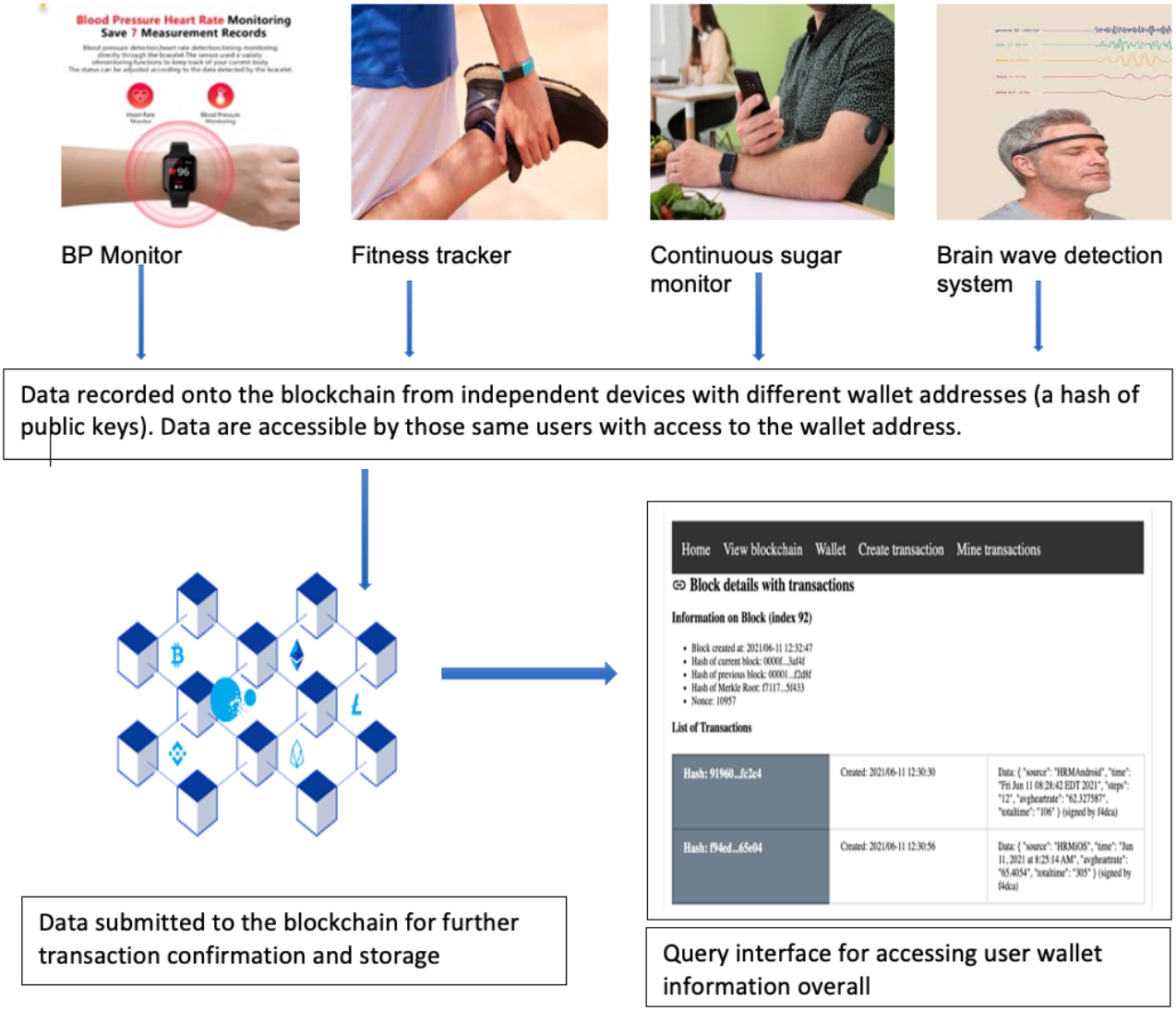
Step 3 of the DSRM: Design of the System

Overview

Figure 1 describes the main use case scenario that motivated our design of the mHealth app. In Figure 1, we see the same individual user using 4 separate devices for monitoring his or her health parameters: blood pressure monitoring device, physical activity (eg, running, jogging, and walking) monitoring, active blood sugar monitoring (eg, for prediabetic or diabetic conditions), and brain wave monitoring. Each device stores and records the individual's personal mHealth data, which can either individually or collectively be used to make inferences about a person's health.

XSL•FO
RenderX

Figure 1. A block diagram showing the process of uploading and managing data with PantherChain. BP: blood pressure.



Use Cases for Storage Retrieval and Messaging

We present different use cases for data storage, data retrieval, and messaging. Apps from different platforms allow users to share data using their private keys to keep the data secure but usable by these apps. These apps can later use whole or part of the user data to provide services, as well as for users themselves to share and control the data provided to the platform.

The use cases are as follows:

1. Use case 1: perform a data collection activity and submit signed data to the blockchain
2. Use case 2: create diverse types of data for the same user from multiple devices using the same private key
3. Use case 3: delink data on the blockchain using private key deactivation (as a blockchain is an immutable structure, data once added to the chain cannot be deleted; however, if the private key is deactivated, that is, delinked from the user, the private components of the data can never be linked back to the user who created it, and any existing encrypted payload data cannot be decrypted in the future or accessed by others)

Other peripheral use cases are as follows:

1. Use case 4: share data across different apps
2. Use case 5: (optionally) sell data in a health care marketplace

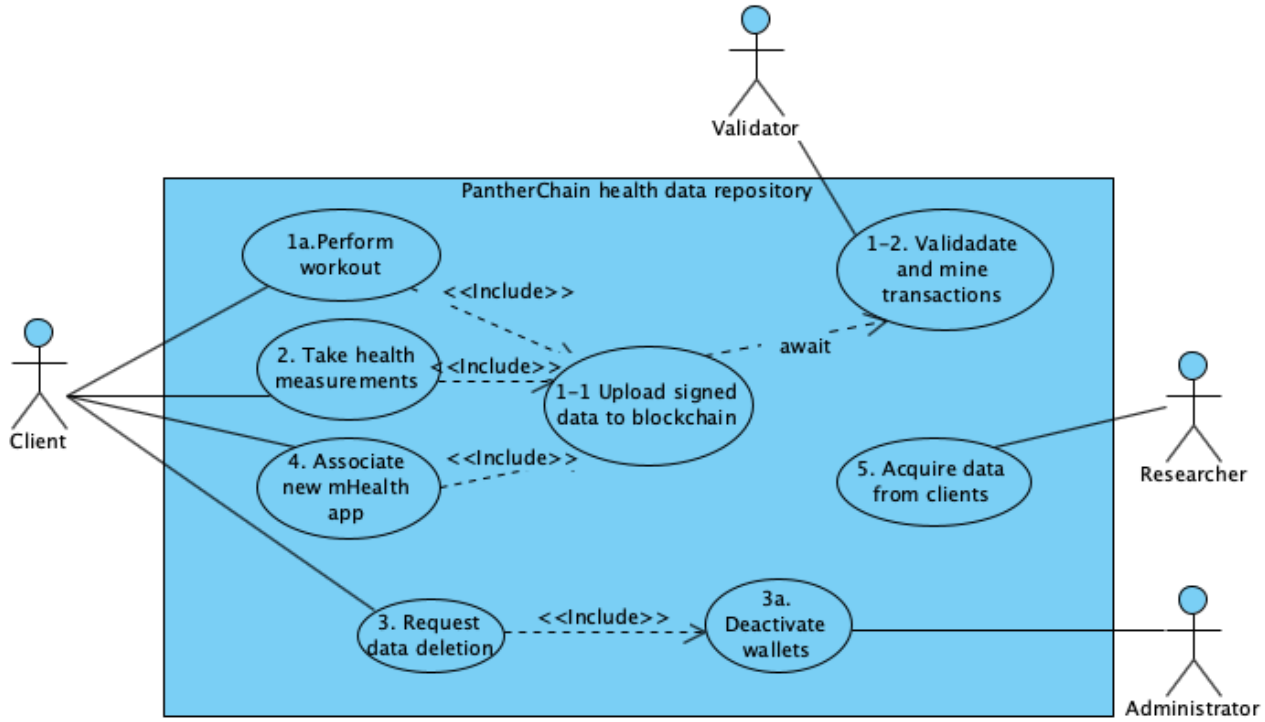
Blockchain design could also incentivize users to create, record, and share authentic, high-quality data and possibly auction the information on health care data marketplaces that access personalized health care data for downstream sale for research. A use case diagram demonstrating all the above use cases and their dependencies is shown in Figure 2.

We extend prior research on configurable blockchain patterns and specifically use the CHASM blockchain pattern and the corresponding instantiated *PantherChain* design to accomplish storage, retrieval, and market functionality for health care data [6]. In the following sections, we describe the design of the PantherChain system and document features of the PantherChain blockchain system used to implement key use cases for pervasive IoT-based mHealth data. The advantage of using the CHASM design pattern is that it is flexible, and each underlying component can be altered to suit the performance, speed, and storage needs of the overall IoT system. Sengupta and Subramanian [6] demonstrated the design and performance functionalities of 4 different blockchains by altering the software

XSL•FO
RenderX

components. We present the architecture of such a system and illustrate the benefits of such an implementation in the following sections, which include high performance and tunability of parameters such as difficulty used in mining, especially when IoT devices and high throughput are required.

Figure 2. Use case model showing all use cases and dependencies. mHealth: mobile health.



The CHASM Design Pattern and PantherChain Implementation

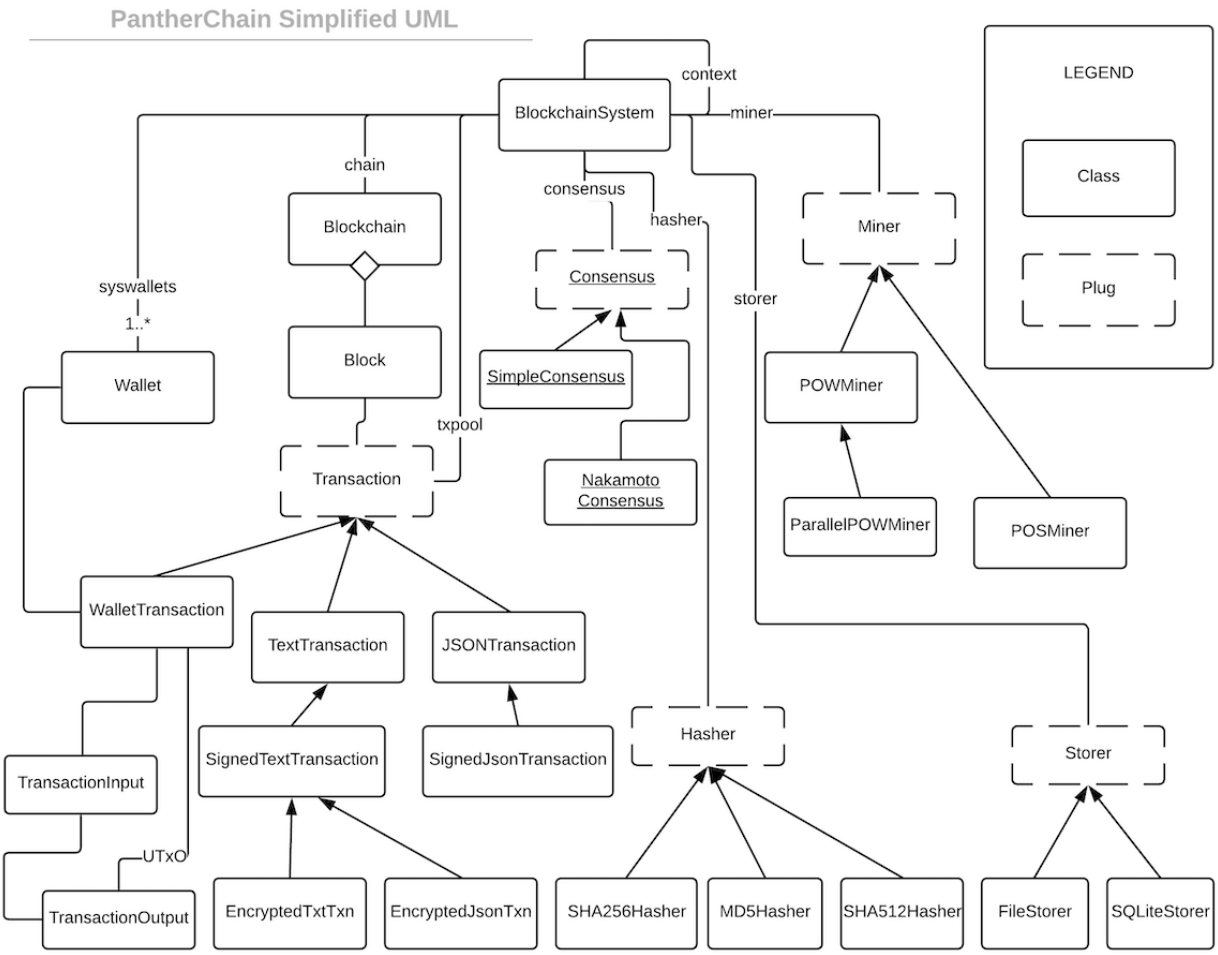
We used the CHASM software design pattern and the implementation of this pattern in Java that we call *PantherChain* [6]. Sengupta and Subramanian [6] described the CHASM in greater detail. However, they did not configure PantherChain or CHASM to support IoT through web service–based application programming interfaces (APIs) for smart contracts. In this implementation, we adapted CHASM's configurable architecture and implemented an IoT-based extension to CHASM in the form of a representational state transfer (REST) API layer that accepts requests from devices and supports high performance and throughput.

In Figure 3, we show the CHASM design pattern by illustrating the classes and their interactions. The CHASM pattern is an extension of the popular context object design pattern [18], a core Java 2 Platform Enterprise Edition pattern that allows different classes to share a single controlling context (interface). In our implementation, the BlockchainSystem class references abstract class implementations of the 4 CHASM components: consensus, hasher, storer, and miner. Henceforth, we will refer to these concepts, implemented using abstract classes, as *plugs*.

In PantherChain, we provided different implementations for each plug. The *hasher* plug has 3 implementations: MD5 (Message Digest Version 5) [19] implemented as MD5Hasher, SHA256 (Secure Hash Algorithm) [20] implemented as SHA256Hasher, and SHA512 [20] implemented as SHA512Hasher. The storer plug has 2 implementations: a serialized JSON file implementation (FileStorer) and a relational implementation using SQLite (SQLiteStorer). The *miner* plug was implemented as a *proof-of-work* (POW) miner with flexible levels of difficulty, similar to Bitcoin [21]. We included 2 implementations of POW: a single-threaded POWMiner and a multi-threaded parallel version (ParallelPOWMiner). Finally, the consensus plug was implemented using a distributed file-sharing mechanism as a proof of concept to demonstrate a distributed Nakamoto [21] consensus, as described by Nakamoto [21] (SimpleConsensus). PantherChain can be extended by replacing the *plugs* (implemented as abstract Java classes). Each plug provides specific abstract methods that are interfaces to the rest of the blockchain system. Finally, PantherChain can use any implementation of transactions and can be used for different types of transactions (eg, text-based or cryptocurrency-based) within the same instance. In PantherChain, both text-based and JSON-based transactions can be signed and encrypted. Interestingly, as the payloads of encrypted JSON transactions are no longer well-formed JSON, the encrypted JSON transaction class is a subclass of signed text transactions (Figure 3).

XSL•FO
RenderX

Figure 3. Unified Modeling Language (UML) class diagram of the PantherChain blockchain.



Step 4 of the DSRM: Implementation of the IoT Interface on the Blockchain

Overview

Although smart contract [22] virtual machines are not implemented within the CHASM pattern, we piggybacked on the Java Virtual Machine combined with an easy-to-use web service–based software as a service system that operates on each PantherChain node as the IoT interface. The current implementation of PantherChain includes a Java 2 Platform Enterprise Edition Servlet–based REST API that allows an external app to use all components of PantherChain. Some of the end points in the PantherChain REST API are as follows: *APIGetKey* (to retrieve an existing wallet or create a new public or private key wallet), APIAddJson (to add a transaction containing JSON data signed with a key), and *APISearchUid* (to search PantherChain based on a unique identification).

To implement a smart contract, the app can make an API connection to a PantherChain node and provide the client with complete app integration. This method allows app developers to use a blockchain in the back end and develop their app logic in the language or platform of their choice. We developed and tested each of our use cases listed above on PantherChain to demonstrate the capabilities of the system and validate it with user-based tests and HIPAA compliance guidelines. We also set forth future research directions for such a model by illustrating incentivization models based on third-party validation of data accuracy, marketplaces for data, etc. The PantherChain implementation also includes the demonstration of Bitcoin-like cryptocurrency (PantherCoin) atop the existing system.

Invocation of the IoT mHealth Data Interface

Integrating PantherChain into an mHealth system is a 2-step process: (1) initializing or retrieving a user wallet (using APIGetKey) and (2) submitting a transaction to the pool using APIAddJson. In our tests, we developed 2 apps, one for an Android watch and the other for an Apple watch, each of which can capture some health data from the sensors, build a JSON data packet, and send it to the API back end. Multimedia Appendices 1 and 2 show sample Android (Java) and iOS (Swift) codes for performing this action.

XSL•FO
RenderX

Results

Step 5 of the DSRM: Demonstration of the Mobile Data Health Care iOS and Android Apps

Overview

To provide a proof of concept of the process, we developed 2 watch apps: one based on Android and the other based on iOS WatchKit. Both were simple apps that were designed to run on a watch. We programmed the Android app to track steps and heart rate data when a *start* button is tapped and submit a signed package containing the average heart rate and total steps taken when the *stop* button is tapped. The iOS WatchKit app worked in the same way, except that, given the iOS watch we used did not have a direct internet connection, the data were uploaded by an iPhone companion app that retrieves the data from the Apple watch and submits a similar signed package to the blockchain. The user can set up both apps with the same wallet (private-public key pair) so that the data from all the devices can be linked to the same user. Alternatively, the user could use a meta-wallet such as metaMask [23] or the Exodus [24] wallet to generate a master public-private key to store independent wallets within the meta-wallet. Overall, irrespective of the approach chosen, all the user's data and confidential information are captured and stored only by a single user. In the following sections, we provide some illustrations of the apps we designed, thus demonstrating the various use cases above.

Implementation and Validation of Use Case 1

In this use case, the personal health care data of users are stored and accessible on the blockchain using a query interface using the private key by a single user.

The first basic use case involves a user performing a basic workout using an Android watch. Figure 4 demonstrates the Android app running and collecting data while the workout is in progress. The app may enable the encryption of the payload for additional security of the data. We will discuss the encryption process further in the *Implementation and Validation of Use Case 3* section. When the workout is stopped, the watch (or the controlling phone if the watch does not have direct internet access) can submit the workout as a new transaction to the blockchain after encrypting it with a secret key (if enabled at the app level).

Figure 4. Android implementation of our app running on an Android smartwatch.

Implementation and Validation of Use Case 2

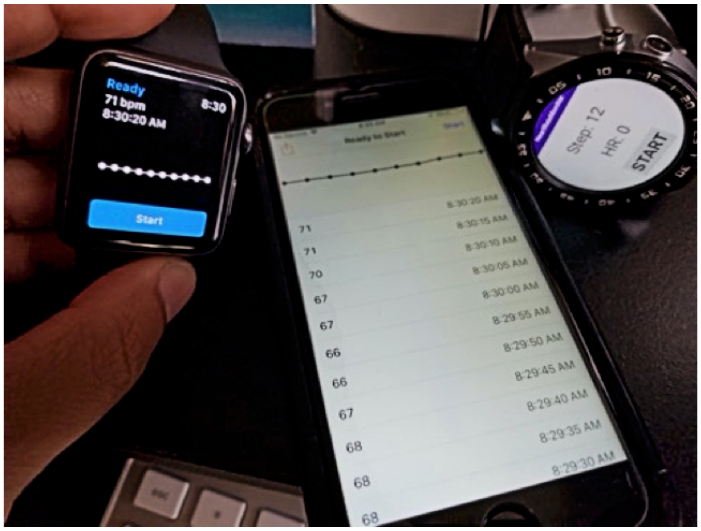
In this use case, personal health care data of diverse types with multiple devices and meta-information are stored on the blockchain.

Although the first use case is the most basic method of tracking mHealth data, similar processes exist in most health monitoring platforms today. The motivation for use case 2 comes from the limitation of the ability to share and aggregate data from different mHealth systems into a single platform. Currently, users do not have the option of using different devices with differing capabilities of collecting health data and reviewing and aggregating all data in a single coherent manner. Our app solves this issue by allowing different apps to collect mHealth data to submit data packages signed with the same wallet to be submitted to the blockchain. Figure 5 shows the 2 apps in iOS WatchKit (left) along with its companion iPhone app (center) and the Android watch app (right), collecting diverse types of data for submission to the same blockchain.

Figure 5 illustrates 2 different devices (an iOS device and an Android device) collecting health care data from the same individual and submitting the data to the blockchain. The user of the app, identified by his private-public key combinations, can access the data within his wallet from a web app (as shown in Figure 1). Such data collected about the user from multiple devices can be used to create a health intelligence dashboard and can also be traded (shared) with data marketplaces.

XSL•FO
RenderX

Figure 5. Different devices collecting health care data.



Implementation and Validation of Use Case 3

In this use case, personal health care data are deleted based on the private key of the individual after deboarding from the system.

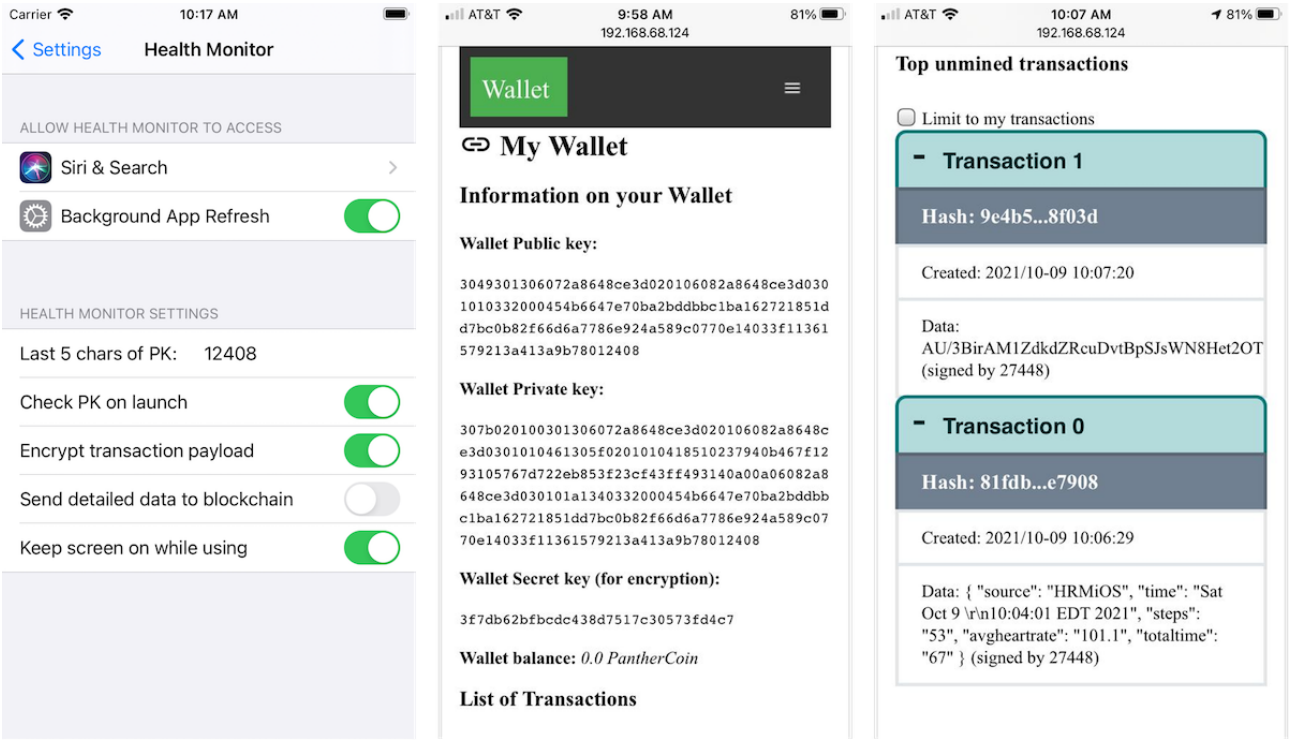
To provide HIPAA compliance, users need to have the ability to delete personal health data from any associated system. As a blockchain is immutable, data cannot be tampered with once it is mined and added to the blockchain (but it can certainly be deleted from the transaction pool before it is mined). However, in our mHealth implementation, HIPAA compliance can be achieved by deactivating the user wallet, which delinks any mined block from the user. If the user chooses to encrypt the data payloads for the transactions, the encrypted components cannot be decrypted without the user's secret key. However, the blockchain integrity can still be maintained with blocks containing the data, although no association with a specific user can be made without the wallet.

We highlight some other potential approaches for privacy implementations in the following sections.

Implementation and Validation of Multi-key Wallets

A wallet in our app can be associated with multiple secret keys for encryption. Deactivation of any key associated with the wallet renders the transactions encrypted with that key inactive. In our implementation, a wallet included a public and private key pair for signing and verification and several advanced encryption standard secret keys for encrypting payloads. Figure 6 demonstrates (1) enabling the user of the mobile app to encrypt transaction payloads; (2) a wallet with a public key, private key, and a single secret key; and (3) a simple signed transaction and a transaction from the same user with an encrypted payload.

Figure 6. Implementation of encrypted transaction payloads in PantherChain and health monitor app.



XSL•FO
RenderX

App Access Restriction

Privacy compliance can be implemented at the app level, where the user requests the deletion of an account, leading to the removal of the user's wallet from the blockchain. Once the user deboards the system, if the private keys are removed from the personal devices of the users, the data can never be reconnected to the user level.

Implementation and Validation of Use Case 4

In this use case, personal health care data with different types of data are aggregated into a blockchain app.

Once data from different apps are uploaded into the blockchain, a third-party verification method can validate the authenticity of the data (this process may differ from blockchain to blockchain but will involve some form of mining strategy). In this proof-of-concept app, we used a POW mining technique (similar to that of Bitcoin but using a lower difficulty level) to combine validated transactions into mined blocks. These validated blocks can be used by apps to aggregate data from every app that uses the same wallet to be associated with the same user. The mining module in the CHASM pattern in PantherChain is implemented as an interface [6]; hence, it can be replaced with other algorithms such as Proof-of-Stake, which can provide deterministic, time-bound confirmations for transactions logged into blocks [25].

Governance Approaches for PantherChain

One of the largest challenges with IoT-based blockchains is related to who governs the blockchain. Governance functions for the blockchain include the ability to control the functionality of the blockchain [26]. This could include altering algorithms for consensus, hashing, storage, and mining. In addition, the blockchain's future functionality enhancements, including the ability to support different apps on top, could be subject to governance. Governance protocols that can be considered for the blockchain can be of many types. Table 2 summarizes the key governance mechanisms of the token.

On the basis of our analysis of an mHealth-based app stack, which interfaces IoT with the blockchain, the *consortium-based* blockchain works best. If a consortium of device manufacturers, app developers, and users cannot be formed, creating a decentralized autonomous organization facilitated by PantherCoin, the inbuilt token of the PantherChain blockchain, will enable such a mechanism. However, the risks of such an approach are greater than having an equitable consortium of stakeholders who invest and control the blockchain and its future enhancements and can ensure decentralization. For a system such as ours that deals with health care, it is recommended that governance not rely on tokenomics; rather, the decentralized functionality of the blockchain should be used overall.

XSL•FO
RenderX

Table 2. Governance mechanisms recommended for the blockchain.

Type of governance	Description	Benefits	Challenges
Consortium-based governance, which is a user group comprising device manufacturers, users, and health app writers; prior apps have included IoT[a]-based power networks [25]	• A group of firms, trusts, and user groups pool in resources and administer the whole network of nodes. • Membership to the consortium would be rule-based and inclusive. • Members could include private organizations, departments of health, and nongovernmental organizations, which are responsible for running and maintaining this structure.	• Such a design automatically ensures that the members of the consortium are invested in governance. • Adoption is almost instantaneous as the network is already seeded by the consortium. • As CHASM[b] is a configurable blockchain, underlying algorithms and methods can easily be altered, as quorum among consortium members is easy to accomplish.	• Typical issues of collusion, exclusion and nonadherence to the rules of governance could slow adoption or reduce the number of users. • The consensus algorithm (and the miner plug) could be moved to proof of stake from proof of work, which is a change at the underlying level. However, with proof-of-work mining, the mining algorithm could be operated by the consortium itself.
DAO[c] [26]	• Governance is typically decided by the voting rights of users who own crypto-tokens generated by the platforms and which are purchased in exchange for fiat currency. • Voting rights are usually proportional to the share of tokens owned by those who choose to govern.	• Everyone ideally has a chance to participate in governance. • It could potentially lead to faster adoption as governance is also incentivized.	• Governance is usually skewed among those token holders who hold the largest number of tokens. • This could result in a dysfunctional blockchain if the tokenomics do not reward users appropriately. • There are risks of rug pulls in the market.
Public blockchain [27]	• This is similar to the creation of any large public blockchain. • All users have equal rights on the network, and peer nodes handle the traffic.	• Public blockchains foster trust among individuals for easier adoption. • Public blockchains are useful for internal purposes as well as external purposes.	• It involves extremely slow adoption. • There are no special incentives to users and node runners. • Network effects will become extremely difficult to run and maintain. • Time taken to modify and roll out changes will disincentivize device manufacturers and users.
Private blockchain controlled by device manufacturers	• This blockchain is similar to hosting a set of nodes and controlling their data privately on the network.	• It is easy to set up as private investors are involved. • Changes to the underlying blockchain are all dictated by individuals on the network overall.	• It is controlled by private actors. • Decentralization of governance, modifications, and data is difficult to accomplish overall.

[a]IoT: internet of things.

[b]CHASM: consensus, hasher, storer, miner.

[c]DAO: decentralized autonomous organization.

Cost of Setting up a Blockchain

There are many types of costs associated with setting up such a blockchain, such as the costs for programming, the cost of hardware setup, the costs for setting up a network, and the cost of app development and integration with device manufacturers. Although this paper will not be sufficient to cover all these costs because of the detailed cost analysis needed, we have documented the cost of hosting our prototype. We hope that designers and developers will be able to interpret the hosting costs from these figures and find these costs significantly lower than those of the public blockchains or other consortium-based blockchains, which require more specialized knowledge and higher bandwidth equipment such as application-specific integrated circuit (ASIC)–based processors for mining.

Our prototype instance was configured and set up on a cloud virtual machine with 1 TB disk space, 128 GB RAM, and Linux with 2 core Intel processors. The cost of setting up 1 instance of the blockchain was US $300 for 3 years. Similarly, nodes could be set up at US $1200 for 3 years. This is for unlimited incoming and outgoing network traffic. We used a common high-performance computing infrastructure cloud available to us from the university. A similar hosting arrangement could be applied to the cloud environments (such as Amazon Web Services, Google, Microsoft, and Heroku) or even with blockchain-based distributed computing clouds such as FileCoin, Storj, or InterPlanetary File System–based systems, which could cost the same amount. On the basis of the number of consortium members and the number of device manufacturers or devices that store data on the cloud, use could also increase or decrease as time goes by. These systems are flexible enough and can

XSL•FO
RenderX

accommodate the need to store only the most recent data (and partial or full replication of data could be supported). As the CHASM architecture is flexible, future versions could include sharding of the data store and other performance enhancements that are typical of large-scale data stores. A full discussion of distributed systems architecture would be out of the scope of this paper as we focus on the privacy and user control of data using an IoT-optimized data store.

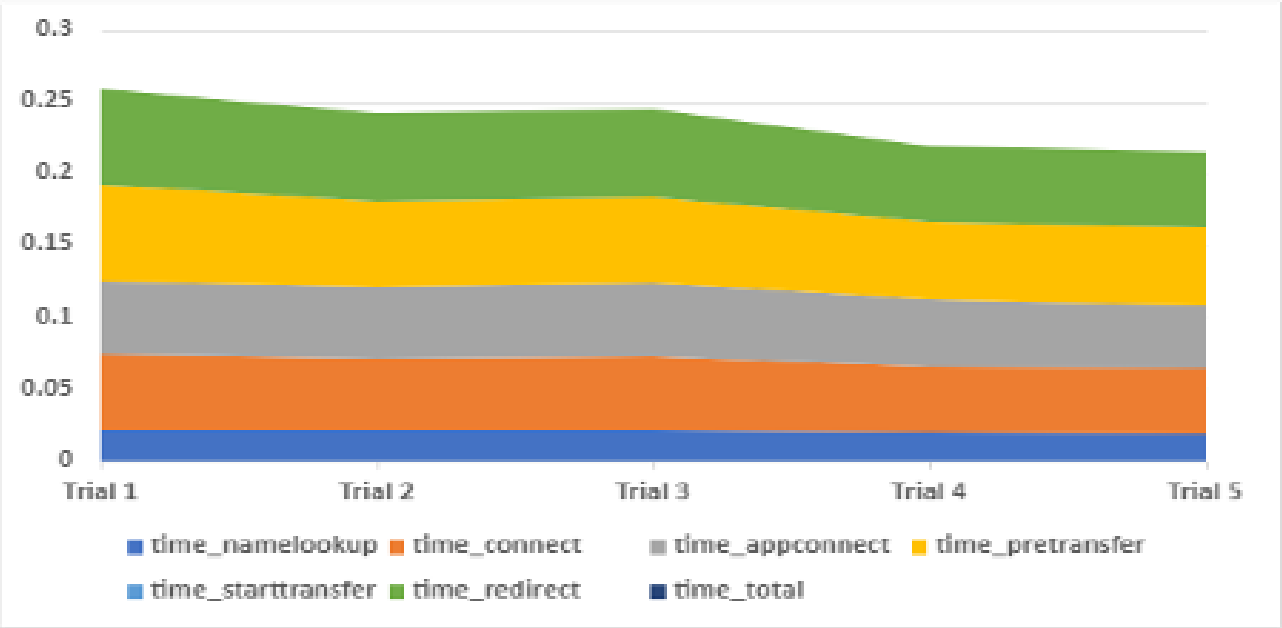
Performance Statistics for the PantherChain System

Although the performance of blockchain systems has been subject to much debate in the literature, it must be noted that performance is measured in terms of the number of transactions that the blockchain can handle per minute overall. For example, the VISA network supports 1700 transactions per second, which indicates a submillisecond access time for any web service request on the network. The Bitcoin and Ethereum networks are much less performant, and these systems have not been benchmarked from an app standpoint.

Our approach to testing the performance of the PantherChain is based on the caveat that the performance of any system can be assessed by measuring the overall times taken to access different components of the system. In our case, owing to the architecture of a distributed and decentralized computing system, we could measure the overall performance of our system in terms of the time taken to access the services of the blockchain. For example, the time taken to access the front end of PantherChain and the time taken to access the REST APIs are discussed above. As we use the Apache Tomcat web server, using the software as a service model, several performance benchmarks of the stack exist. Notably, the Center for Internet Security benchmarks for Windows and Linux operating systems and several processor-related benchmarks already exist [27]. Similarly, Apache Tomcat and Java performance benchmarks for a variety of hardware and software, including combinations of Java Development Kit and Java Virtual Machine implementations, have already been performed [28]. For the performance testing of different configurations of blockchains, please refer to the study by Sengupta and Subramanian [6].

We benchmarked the standard typical use cases of the front end of the PantherChain system and the basic API for obtaining the wallet key after generating it. In Figure 7, we plot the means of the API response times on 5 dimensions commonly used to measure the performance of the getKey and uid APIs and the PantherChain user interface.

Figure 7. Mean performance of application programming interface response times from PantherChain implementation.



Validation of the Artifact

We provide a complete validation of our process, against the requirements, including the HIPAA compliance checklist for mobile devices [29], in Textbox 3.

XSL•FO
RenderX

Textbox 3. Confirmation of the achievements of the design goals for the system.

Health Insurance Portability and Accountability Act guidelines and how our internet of things mobile health app and blockchain design support the functionalities

- Use a password or other user authentication
  - Users use a combination of private and public keys that they must access for personal mobile health data.
- Install and enable encryption
  - All data are encrypted at multiple levels. At the storage level, users could choose to encrypt and send data using their public key (or a separate key), and their wallet software could enable access to these data later.
- Install and activate remote wiping or remote disabling
  - As discussed, the deletion system for personalized mobile health data can be deleted at the app level and the individual level.
- Disable and do not install or use file-sharing apps
  - File or data sharing is not enabled at the app layer, although the blockchain resides on a decentralized network. The access to data is provided only to those with the private and corresponding public key (wallet) and not to anyone else.
- Install and enable a firewall
  - The strong encryption provided by the network and the iOS or Android operating systems for mobile health data provides necessary protections. Overall, if firms choose to provide an operating system–level firewall for other types of data from devices such as blood pressure monitors, they can do the needful.
- Install and enable security software
  - This is outside the scope of our app, although we believe that pervasive devices are compliant with security.
- Keep your security software up to date
  - Blockchain software, when updated, will reflect the same at the app layer, and users will be able to directly contact the apps.
- Research mobile apps before downloading
  - This is a user characteristic, and users who work on data should be cautious while downloading apps. Either way, without private and public keys, users will not be able to move data onto the blockchain.
- Maintain physical control
  - Pervasive health care data are strongly controlled by the user's ability to protect their public and private keys and, therefore, control access.
- Use adequate security to send or receive health information over public Wi-Fi networks
  - Data can be encrypted end to end using HTTPS protocols to allow only secure apps from devices to communicate with the blockchain's web service layer.
- Delete all stored health information before discarding or reusing the mobile device
  - Once a request is made to delete a particular private key's data by the user, the user can deactivate all the data from the app by preventing access to it forever. Whenever a mobile device needs to be reused, it is up to the user to remove his or her private and public keys from the device.

Discussion

Evaluation of the Prototype

In this study, we developed a complete blockchain-based mHealth data collection system using PantherChain, an implementation of the CHASM-based blockchain framework. We developed mobile apps capable of running on smartwatches to collect personal health data, which are signed with the user's private keys, and demonstrated that this process could be implemented in high-performance apps. We tested our results with a proof of concept, as explained in detail in previous sections, using both an iOS and Android mobile app and the blockchain. We can clearly show evidence for access to diverse types of health information [6] by the same user.

Further results demonstrating the performance of our app in the context of health parameters and different operating system versions are documented in the following paragraphs.

We collected performance data by running our proof-of-concept apps on different devices and simulators. The proof-of-concept iOS WatchKit app was tested on the following platforms:

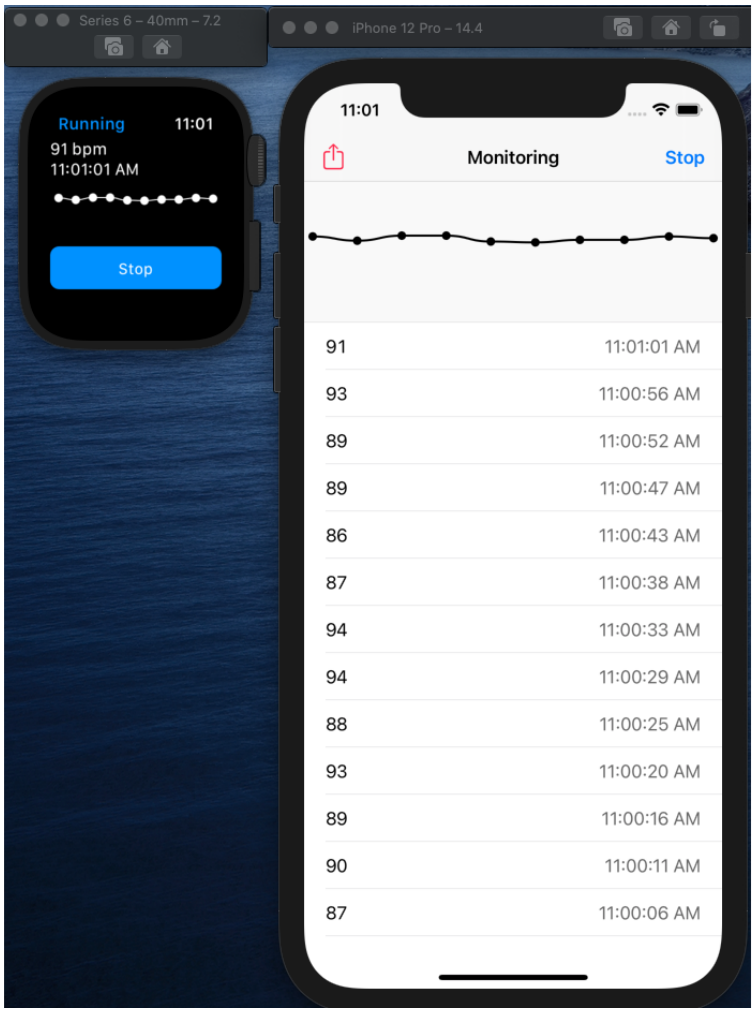
XSL•FO
RenderX

1. iPhone 12 Pro simulator running iOS 14.4 paired with Apple Watch Series 6 40 mm simulator running WatchOS 7.2 (Figure 8).
2. iPhone 6 device running iOS 12.5.4 paired with Apple Watch (original) running WatchOS 4.3.2 (Figure 5)
3. iPhone 12 device running iOS 14.6 paired with Apple Watch Series 3 running WatchOS 7.3.2

Our proof-of-concept Android watch app was tested on the following platforms:

1. Android Watch device ZGPAX S99c running Android 5.1 Lollipop (Figure 4)
2. Android Watch simulator (not reported in results because of unavailability of health simulators) running Wear OS 9.0 (Android Pie)

Figure 8. Depicting the simulators showing Android apps sending data to the blockchain.



The host platform for the 2 simulators we used was a MacBook Pro 2018 (Apple Inc, 2.6 GHz Core i7, 32 GB RAM) running macOS Catalina. The work network was a wired 1 Gbps network, and the home network had a maximum upload speed of 5 Mbps. Note that the Android watch device we used had only a 2.4 GHz wireless adapter (which is typical of most low-cost, high-performance IoT platforms, as was available during the writing of this paper). We tested on some old hardware to ensure that our apps could run with adequate performance on low-performance IoT devices. After running our apps, which were configured to send summary as well as detailed workout data, and varying the time of workouts, we noticed that the performance of submitting workout data to PantherChain primarily depended on network latency, network throughput, and HTTP protocol handshake time. The round-trip time from device to PantherChain was observed to be as low as 28 milliseconds (simulator in 1 Gbps network) to a maximum of 724 milliseconds (Android 5.1 device on a 2.4 GHz home wireless network). We collected data for workouts ranging from 1-minute to 30-minute durations, and even with detailed data uploads, the performance was reasonable and as expected for any network-connected device. We plot the graphs of the time taken to submit the data to the blockchain in milliseconds (y-axis) versus the size of the data (in bytes) payload on the corresponding device for the 4 configurations in Figure 9.

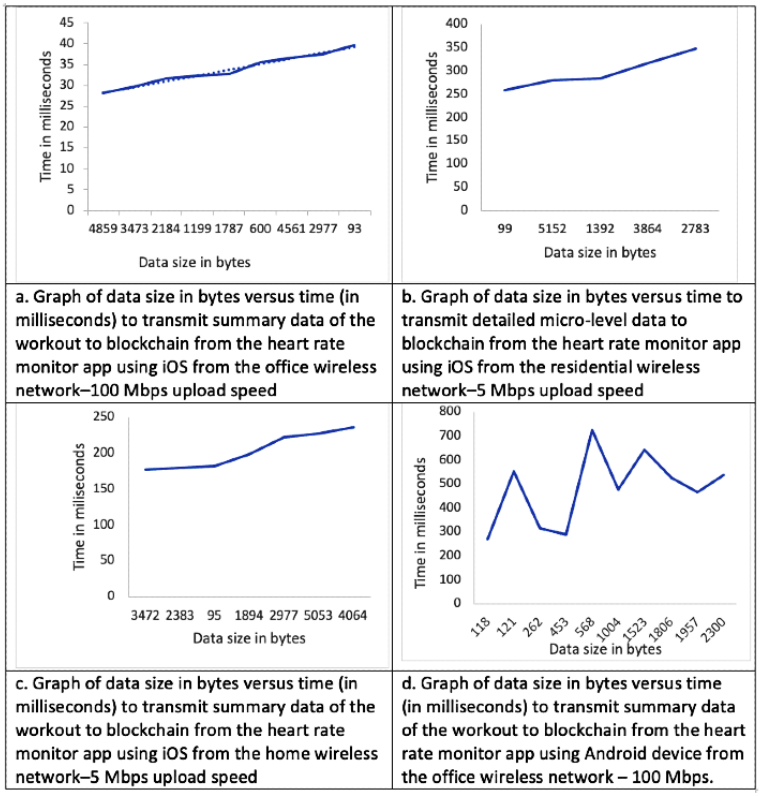
The 4 configurations correspond to using the heart rate monitor app running in 4 different modes on different network configurations, as shown in Table 3.

The results demonstrate that our platform and process for collecting and submitting health data to a blockchain does not degrade user experience in terms of performance. Please note that the performance times reported are only based on submission to the blockchain and do not include mining performance, which is not necessary unless the user wants to have the data committed to the blockchain for potential

XSL•FO
RenderX

marketplace use. Mining performance comparisons for various difficulty levels (by varying the mining algorithm) of PantherChain are available from the study by Sengupta and Subramanian [6].

Figure 9. Blockchain submission modes for internet of things performance testing for different network and device combinations for the Heart Rate Monitor (HRM) app.



a. Graph of data size in bytes versus time (in milliseconds) to transmit summary data of the workout to blockchain from the heart rate monitor app using iOS from the office wireless network–100 Mbps upload speed

b. Graph of data size in bytes versus time to transmit detailed micro-level data to blockchain from the heart rate monitor app using iOS from the residential wireless network–5 Mbps upload speed

c. Graph of data size in bytes versus time (in milliseconds) to transmit summary data of the workout to blockchain from the heart rate monitor app using iOS from the home wireless network–5 Mbps upload speed

d. Graph of data size in bytes versus time (in milliseconds) to transmit summary data of the workout to blockchain from the heart rate monitor app using Android device from the office wireless network – 100 Mbps.

Table 3. The 4 configurations and the type of data collected.

Device	Type of data collected	Network configuration
iOS	Summary	Work; wired
iOS	Summary	Home; wireless
iOS	Detail	Home; wireless
Android	Detail	Home; wireless

Comparison With Prior Work

Although some prior works have provided a methodology for achieving tamper resistance and privacy in storing mHealth data in blockchain (eg, Ichikawa et al [8]), our implementation takes it further in a full blockchain implementation using actual mHealth data apps. Fang et al [17] suggested using a permissioned private blockchain to accomplish some of the GDPR capabilities to enable users to control their data. Scholars have developed and demonstrated apps for blockchain in the areas of health care exchanges, mobile data exchanges, and centralized health record storage. Such work has mostly been with enterprise health care systems such as electronic health care records and whose primary use case for the blockchain is for decentralizing data storage [30-34]. Although such apps are built and deployed on either Hyperledger or Ethereum frameworks, they are not necessarily tested for IoT compliance, which is at the other end of the performance, data size, and speed spectrum. Griggs et al [35] discussed that with public blockchains, no user data should be stored within smart contracts as HIPAA compliance indicates that such data are not private and are accessible to all. Our configurable blockchain design considers this limitation and solves this problem by encrypting the data and allowing data access only through wallets that are owned by users. Similarly, we can also configure the entire blockchain on a private network and not on public blockchain infrastructure, which prevents public viewing of such data.

Several prior papers have cited and used Ethereum as an underpinning technology for smart contract–based blockchain designs. However, with Ethereum as the blockchain, on-chain storage is expensive, and for IoT apps such as edge devices, which require storage of data on the chain for fast retrieval or health care analytics, Ethereum may not be best suited as a blockchain. A recent estimate for the Ethereum blockchain showed that storing 1 MB of data on chain would cost approximately US $76,000 [36]. As a result, on-chain storage on Ethereum is not recommended for current apps. Storing off chain will also be significantly inefficient because of the gas costs on Ethereum for writing a transaction and operating a

XSL•FO
RenderX

smart contract, which today costs approximately US 25$ to US 60$ per transaction, which would be more expensive than some of the low-end smartwatches in the market.

Another limitation of prior work is that with mHealth data and pervasive computing, the need for data platforms to operate in the context of multi-platform data sources has not been tested. Pervasive (edge) devices that need to transmit data frequently to back-end data stores present specific network transmission time (performance), security, and privacy challenges. The web services architecture (CHASM) that we deploy is platform independent, and we demonstrate using Android and iOS apps with different network configurations, all communicating with the same blockchain web service for 1 individual user. Overall, we believe that our approach is a novel method for designing and developing and demonstrating the functionalities of a HIPAA mobile data–compliant system that comprises pervasive mobile data apps on multiple devices and a custom configurable blockchain that is platform independent and performant with respect to IoT characteristics. We conclude our research by stating the benefits of using such a configurable blockchain architecture (CHASM) with a web service stack that receives data from IoT devices.

Limitations

Our proof-of-concept solution has been tested for scalability with 4 nodes and works on a secure public network. We tested fail-safeness and replication of data across nodes for the ascertained (mined) data and blockchain. In the future, the architecture can be extended to include a larger number of nodes, and distributed data replication can be tested. However, we are confident that the CHASM architecture will support such scalability as the underlying network and data distribution interfaces are pluggable. Similarly, we believe that our tests performed with >5 devices, including simulators and health apps that track physical activity, can be scaled to a larger number of devices and several types of apps. The PantherChain implementation is device agnostic, and data stored on the blockchain are agnostic to the source of data.

The heterogeneous usability of the blockchain used by providers and device manufacturers, who would now lose control over the data (or would have to purchase data from their customers), to analyze them will lead to reduced times in designing and developing more efficient data. The need for users to share revenues and profits with other users who produce data will be difficult. The risk of losing the private key to access one's data will make the data completely unusable forever. Therefore, the storage of the private key (or private key generation algorithms) needs to be adhered to closely.

Conclusions

Overall, we demonstrate the benefits of using blockchain in presenting a unique and novel architecture for mHealth data using multiple devices. Our implementation of a unique blockchain that is configurable with respect to its subcomponents and API layer provides the necessary flexibility and security in addressing privacy, security, and data ownership issues with personal health data. Our findings from implementing the proof of concept with multiple devices and personal data are as follows:

1. With IoT devices and personal mHealth data, blockchains can provide the flexibility of storing, retrieving, and accessing individual user data, despite multiple devices and operating systems generating such data about the individual.
2. Aggregation of data on the individual and providing only individual access to blockchain data using encryption protects the user's privacy. Using the wallet approach, the user can also transfer or send a copy of his or her data to other wallets through a separate transaction.
3. Our design of mHealth apps on different devices and operating system agnostic apps, which communicate with the blockchain through web services, supports IoT throughput requirements.
4. As stated in the test results above, we demonstrate the performance compliance for IoT devices as well as HIPAA compliance for user data with our design.

Using this implementation, in the future, we could create a *data marketplace* for personalized health care data, where anonymous users can control the generation, quality, and monetization of their health care data, instead of giving data away for free.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Invocation of mobile health data uploaded to PantherChain (Android and Java).
[DOCX File , 13 KB - mhealth_v10i1e32104_app1.docx]

Multimedia Appendix 2
Invocation of mobile health data uploaded to PantherChain (iOS and Swift).
[DOCX File , 14 KB - mhealth_v10i1e32104_app2.docx]

References

1. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng 2018 Mar 19;2(3):158-164. [doi: 10.1038/s41551-018-0195-0] [Medline: 31015713]

XSL•FO
RenderX

2. Esmaeilzadeh P, Mirzaei T. The potential of blockchain technology for health information exchange: experimental study from patients' perspectives. J Med Internet Res 2019 Jun 20;21(6):e14184 [FREE Full text] [doi: 10.2196/14184] [Medline: 31223119]
3. Nahra KJ. HIPAA security enforcement is here. IEEE Secur Privacy Mag 2008 Nov;6(6):70-72. [doi: 10.1109/msp.2008.143]
4. Kempfert A, Reed B. Health care reform in the United States: HITECH Act and HIPAA privacy, security, and enforcement issues. FDCC Quarterly 2011;61(3):240-273.
5. Subramanian H. Decentralized blockchain-based electronic marketplaces. Commun ACM 2017 Dec 27;61(1):78-84. [doi: 10.1145/3158333]
6. Sengupta A, Subramanian H. CHASM: A blockchain design pattern. In: Proceedings of the International Conference on Information Systems (ICIS 2020).: Association of Information Systems; 2020 Presented at: International Conference on Information Systems (ICIS 2020); December 14, 2020; Hyderabad, India URL: https://tinyurl.com/2p9882vt
7. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inform Syst 2014 Dec 08;24(3):45-77. [doi: 10.2753/mis0742-1222240302]
8. Ichikawa D, Kashiyama M, Ueno T. Tamper-resistant mobile health using blockchain technology. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e111 [FREE Full text] [doi: 10.2196/mhealth.7938] [Medline: 28747296]
9. Park YR, Lee E, Na W, Park S, Lee Y, Lee J. Is blockchain technology suitable for managing personal health records? Mixed-methods study to test feasibility. J Med Internet Res 2019 Feb 08;21(2):e12533 [FREE Full text] [doi: 10.2196/12533] [Medline: 30735142]
10. Beinke JH, Fitte C, Teuteberg F. Towards a stakeholder-oriented blockchain-based architecture for electronic health records: design science research study. J Med Internet Res 2019 Oct 07;21(10):e13585 [FREE Full text] [doi: 10.2196/13585] [Medline: 31593548]
11. El-Gazzar R, Stendal K. Blockchain in health care: hope or hype? J Med Internet Res 2020 Jul 10;22(7):e17199 [FREE Full text] [doi: 10.2196/17199] [Medline: 32673219]
12. Hylock RH, Zeng X. A blockchain framework for patient-centered health records and exchange (HealthChain): evaluation and proof-of-concept study. J Med Internet Res 2019 Aug 31;21(8):e13592 [FREE Full text] [doi: 10.2196/13592] [Medline: 31471959]
13. Huberty J, Puzia ME, Larkey L, Irwin MR, Vranceanu A. Use of the consumer-based meditation app calm for sleep disturbances: cross-sectional survey study. JMIR Form Res 2020 Nov 13;4(11):e19508 [FREE Full text] [doi: 10.2196/19508] [Medline: 33185552]
14. Tiron R, Lyon G, Kilroy H, Osman A, Kelly N, O'Mahony N, et al. Screening for obstructive sleep apnea with novel hybrid acoustic smartphone app technology. J Thorac Dis 2020 Aug;12(8):4476-4495 [FREE Full text] [doi: 10.21037/jtd-20-804] [Medline: 32944361]
15. Huang Y, Zhang J, Duan J, Xiao B, Ye F, Yang Y. Resource allocation and consensus on edge blockchain in pervasive edge computing environments. In: Proceedings of the IEEE 39th International Conference on Distributed Computing Systems (ICDCS).: IEEE; 2019 Presented at: IEEE 39th International Conference on Distributed Computing Systems (ICDCS); July 7-10, 2019; Dallas, TX, USA URL: https://ieeexplore.ieee.org/document/8884875 [doi: 10.1109/icdcs.2019.00147]
16. Subramanian H, Liu R. Blockchain and Smart Contract: A Review. Journal of Database Management 2021;32:vii-xxvi.
17. Fang HS, Tan TH, Tan YF, Tan CJ. Blockchain personal health records: systematic review. J Med Internet Res 2021 Apr 13;23(4):e25094 [FREE Full text] [doi: 10.2196/25094] [Medline: 33847591]
18. Alur D, Crupi J, Malks D. Core J2EE Patterns: Best Practices and Design Strategies. New Jersey, USA: Prentice Hall; 2003.
19. Rivest R. RFC 1321 - The MD5 Message-Digest Algorithm. 1992. URL: https://www.rfc-editor.org/info/rfc1321 [accessed 2022-01-04]
20. Lilly G. Device for and method of one-way cryptographic hashing. Google Patents. 2004. URL: https://patents.google.com/patent/US6829355B2/en [accessed 2022-01-04]
21. Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. Decentralized Business Review. 2008. URL: https://bitcoin.org/bitcoin.pdf [accessed 2022-01-04]
22. Benahmed S, Pidikseev I, Hussain R, Lee J, Kazmi S, Oracevic A. A comparative analysis of distributed ledger technologies for smart contract development. In: Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC).: IEEE; 2019 Presented at: IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC); Sept. 8-11, 2019; Istanbul, Turkey. [doi: 10.1109/pimrc.2019.8904256]
23. MetaMask. URL: https://metamask.io/ [accessed 2022-01-12]
24. Exodus. URL: https://exodus.io [accessed 2022-01-12]
25. Bentov I, Lee C, Mizrahi A, Rosenfeld M. Proof of activity: extending bitcoin's proof of work via proof of stake [Extended Abstract]y. SIGMETRICS Perform Eval Rev 2014 Dec 08;42(3):34-37. [doi: 10.1145/2695533.2695545]
26. Lumineau F, Wang W, Schilke O. Blockchain governance—A new way of organizing collaborations? Organ Sci 2021 Mar;32(2):500-521. [doi: 10.1287/orsc.2020.1379]

XSL•FO
RenderX

27. Zhen MJ, Hassan A, Hamann G, Flora P. Automated performance analysis of load tests. In: Proceedings of the IEEE International Conference on Software Maintenance. Canada: IEEE; 2009 Presented at: IEEE International Conference on Software Maintenance; Sept 2009; Edmonton, Canada. [doi: 10.1109/icsm.2009.5306331]
28. Official phoronix test suite tests. Openbenchmarking. 2021. URL: https://openbenchmarking.org/tests [accessed 2022-01-04]
29. How can you protect and secure health information when using a mobile device? HealthIT. 2013. URL: https://archive.healthit.gov/providers-professionals/how-can-you-protect-and-secure-health-information-when-using-mobile-device [accessed 2022-01-04]
30. Alshehri S, Radziszowski S, Raj R. Secure access for healthcare data in the cloud using ciphertext-policy attribute-based encryption. In: Proceedings of the IEEE 28th International Conference on Data Engineering Workshops. 2012 Presented at: IEEE 28th International Conference on Data Engineering Workshops; April 1-5, 2012; Arlington, VA, USA p. 143-146. [doi: 10.1109/icdew.2012.68]
31. Cao S, Zhang G, Liu P, Zhang X, Neri F. Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain. Inform Sci 2019 Jun;485:427-440. [doi: 10.1016/j.ins.2019.02.038]
32. Dubovitskaya A, Baig F, Xu Z, Shukla R, Zambani PS, Swaminathan A, et al. ACTION-EHR: Patient-centric blockchain-based electronic health record data management for cancer care. J Med Internet Res 2020 Aug 21;22(8):e13598 [FREE Full text] [doi: 10.2196/13598] [Medline: 32821064]
33. Girardi F, De Gennaro G, Colizzi L, Convertini N. Improving the healthcare effectiveness: the possible role of EHR, IoMT and Blockchain. Electronics 2020 May 26;9(6):884. [doi: 10.3390/electronics9060884]
34. Magyar G. Blockchain: solving the privacy and research availability tradeoff for EHR data: a new disruptive technology in health data management. In: Proceedings of the IEEE 30th Neumann Colloquium (NC).: IEEE; 2017 Presented at: IEEE 30th Neumann Colloquium (NC); Nov. 24-25, 2017; Budapest, Hungary p. 135-140. [doi: 10.1109/nc.2017.8263269]
35. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T. Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J Med Syst 2018 Jun 06;42(130):1-7. [doi: 10.1007/s10916-018-0982-x] [Medline: 29876661]
36. What is the cost to store 1KB, 10KB, 100KB worth of data into the ethereum blockchain?. 2016. URL: https://ethereum.stackexchange.com/questions/872/what-is-the-cost-to-store-1kb-10kb-100kb-worth-of-data-into-the-ethereum-block [accessed 2022-01-04]

Abbreviations

API: application programming interface
ASIC: application-specific integrated circuit
CHASM: consensus, hasher, storer, miner
DSRM: Design Science Research Methodology
GDPR: General Data Protection Regulation
HIPAA: Health Insurance Portability and Accountability Act
IoT: internet of things
MD5: Message Digest Version 5
mHealth: mobile health
POW: proof-of-work
REST: representational state transfer
SHA256: Secure Hash Algorithm

Edited by L Buis; submitted 14.07.21; peer-reviewed by A Carvalho, S Rostam Niakan Kalhori; comments to author 01.09.21; revised version received 26.10.21; accepted 10.12.21; published 20.01.22.

Please cite as:
Sengupta A, Subramanian H
User Control of Personal mHealth Data Using a Mobile Blockchain App: Design Science Perspective
JMIR Mhealth Uhealth 2022;10(1):e32104
URL: https://mhealth.jmir.org/2022/1/e32104
doi:10.2196/32104
PMID:35049504

©Arijit Sengupta, Hemang Subramanian. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 20.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Multipurpose Mobile Apps for Mental Health in Chinese App Stores: Content Analysis and Quality Evaluation

Xiaoqian Wu[1,2], MS; Lin Xu[1,2], MS; PengFei Li[3], PhD; TingTing Tang[4], MS; Cheng Huang[1,2], MD

[1]College of Medical Informatics, Chongqing Medical University, Chongqing, China
[2]Medical Data Science Academy, Chongqing Medical University, Chongqing, China
[3]School of Public Health, Weifang Medical University, Weifang, China
[4]The Children's Hospital of Chongqing Medical University, Chongqing, China

Corresponding Author:
Cheng Huang, MD
College of Medical Informatics
Chongqing Medical University
No.1 Yixueyuan Road, Yuzhong District
Chongqing, 400016
China
Phone: 86 023 6848 0060
Email: huangcheng@cqmu.edu.cn

Abstract

Background: Mental disorders impose varying degrees of burden on patients and their surroundings. However, people are reluctant to take the initiative to seek mental health services because of the uneven distribution of resources and stigmatization. Thus, mobile apps are considered an effective way to eliminate these obstacles and improve mental health awareness.

Objective: This study aims to evaluate the quality, function, privacy measures, and evidence-based and professional background of multipurpose mental health apps in Chinese commercial app stores.

Methods: A systematic search was conducted on iOS and Android platforms in China to identify multipurpose mental health apps. Two independent reviewers evaluated the identified mobile apps using the Mobile App Rating Scale (MARS). Each app was downloaded, and the general characteristics, privacy and security measures, development background, and functional characteristics of each app were evaluated.

Results: A total of 40 apps were analyzed, of which 35 (87.5%) were developed by companies and 33 (82.5%) provided links to access the privacy policy; 21 (52.5%) apps did not mention the involvement of relevant professionals or the guidance of a scientific basis in the app development process. The main built-in functions of these apps include psychological education (38/40, 95%), self-assessment (34/40, 85%), and counseling (33/40, 82.5%). The overall quality average MARS score of the 40 apps was 3.54 (SD 0.39), and the total score was between 2.96 and 4.30. The total MARS score was significantly positively correlated with the scores of each subscale (r=0.62-0.88, $P<.001$). However, the user score of the app market was not significantly correlated with the total MARS score (r=0.17, $P=.33$).

Conclusions: The quality of multipurpose mental health apps in China's main app market is generally good. However, health professionals are less involved in the development of these apps, and the privacy protection policy of the apps also needs to be described in more detail. This study provides a reference for the development of multipurpose mental health apps.

(JMIR Mhealth Uhealth 2022;10(1):e34054) doi:10.2196/34054

KEYWORDS

mobile apps; app; mental health; mHealth; content analysis

Introduction

Anxiety, depression, stress, and other mental health conditions are increasing worldwide, involving family, study, work, social intercourse, and other aspects. Nearly 1 billion people worldwide have been found to have mental disorders [1]. In China, the weighted lifetime prevalence of mental diseases, except dementia, among the population over the age of 18 years is as high as 16.57% [2]. Although mental disorders will impose varying degrees of burden on patients and their surroundings,

XSL•FO
RenderX

only 15.7% of individuals with lifelong mental disorders in China seek help [3]. This situation is still obvious in cities with high economic development. A study in Shanghai, China, revealed that approximately 21.4% of the subjects reported depressive symptoms but only 4.7% sought mental health services [4]. The general underutilization of mental health services is worrying. The main reasons for this are the shortage of resources, the limited number of mental service professionals, and the uneven geographical distribution of mental health services in China [5]. Additionally, stigmatization of mental health services [6], low perceived demand [7], and economic constraints make people reluctant to take the initiative to accept mental health services [8].

Mobile apps are considered an effective way to eliminate these obstacles, improve mental health awareness [9], and contribute to symptom tracking and self-management [10]. Mobile apps are not limited to a particular time and place and realize the large-scale provision of cost-effective medical services, especially for people who find it difficult to receive traditional psychological services, for example, the population in rural areas with relatively low economic development and those who find it difficult to receive services face-to-face because of special reasons. During the COVID-19 outbreak, the spread and uncertainty of the pandemic caused a pessimistic mood of anxiety and fear for some patients, medical workers, and the general public [11]. This led to a sudden increase in mental health problems and their higher incidence rate, and mental health needs worldwide [12,13]. However, social and interpersonal networks were relatively closed because of epidemic prevention and control. Thus, traditional mental health services were difficult to obtain. Mental health apps can overcome the limitation of distance and expand the scope of psychological counseling for people. This also highlights the potential of digital health in improving the coverage of mental health services [14,15].

The advantages of mobile apps and the growth of mental health demand are making mental health and adaptive mobile health (mHealth) apps increasingly popular. Furthermore, there is an urgent need for more research to promote the formulation of better mental health service recommendations, especially in China's huge untapped market [16]. The main categories of mental health apps are assessment, tracking or monitoring, treatment, and multipurpose [17]. However, the multipurpose mental health apps integrate evaluation, monitoring, treatment, and other mental health services into 1 platform to provide users with one-stop services. They are the most popular apps for all ages [17]. However, there is currently no specific evaluation for multipurpose mental health apps. One study referred to multipurpose mental health apps, but there is a lack of standardized measures to evaluate and compare the quality of apps [18]. A study searched and evaluated China's mental health apps [19]. However, some features of the apps, such as the professional background of app development, the theme distribution of the function, and the user privacy protection policy are still unclear. The privacy protection of applications is an important reference for people to choose mental health apps [20], and the lack of a professional background in the process of app development may reduce users' confidence in the apps. Simultaneously, we found that great changes have taken place in China's app market with the development of the internet and cell phone manufacturers. The app market developed by cell phone manufacturers replaced third-party app markets, such as 360 and Baidu, and occupied the main share of China's app market, together with Tencent My App [21].

Therefore, this study aims to investigate the characteristics of multipurpose mental health apps in China, evaluate their quality, and describe the main functions, user privacy protection, and professional background of the development process of multipurpose mental health apps in the current market in order to help users make more informed choices and provide reliable evidence for app developers.

Methods

Systematic Search Strategy

This study featured a systematic search and content analysis of multipurpose apps on mental health in Chinese app stores on December 17, 2020. Huawei cell phones occupy the first place in the smartphone market share in mainland China [22], and the Huawei AppGallery has become the largest Android app store in China [21]. Moreover, Tencent My App, provided by the Chinese internet giant Tencent, is the second-largest Android app store in China [21]. Thus, we searched the Apple App Store (for iOS apps), Tencent My App, and Huawei AppGallery (for Android apps).

By preliminary test searches, the following keywords were determined: psychology, psychological counseling, psychological intervention, emotion, stress relief, anxiety, and depression. These keywords were searched anonymously using Chinese language terms in the app stores not logged into any user accounts. All search results were collected to ensure that all potentially relevant apps were captured. If an app exists in both iOS and Android and has the same design and content, the Android version was evaluated.

Eligibility Criteria

After removing duplicates, each potentially suitable app was reviewed by 2 independent researchers based on the app name, screenshots, and description. In this round, apps were included if they (1) provide multiple mental health services, (2) focus on individual mental health consumers seeking professional help, and (3) are in the Chinese language. Apps were excluded if they (1) focus on content unrelated to mental health services, such as social and communication apps, e-books, heartbeat measurement, and pulse measurement; (2) target mental health service providers, such as doctors, nurses, or counselors; and (3) provide a single function only. All apps that met the inclusion criteria were downloaded onto test devices. Of these, apps were excluded if they are not usable because of technical errors or require special authentication (such as an enterprise or a school).

Data Extraction

The relevant information provided by the app market was extracted to evaluate the descriptive features of the apps. The general characteristics of the apps, including platform,

XSL•FO
RenderX

developer, target user, update time, star rating, and downloads, were recorded. Additionally, combined with the researchers' use of the apps, the characteristics of the apps with regard to personal privacy protection and professional development background were extracted (Table 1). Furthermore, through a literature review and group discussion, we divided the main services provided by mental health apps in China into 6 categories: psychoeducation, counseling, self-assessment, question-and-answer (Q&A) module, stress relief, self-monitoring, and management. The characteristics of the apps and the main services provided were recorded by 2 independent researchers. All differences were resolved through discussion until the researchers agreed upon the results.

Table 1. General information collected for each app.

Assessment measure	Definition and values
Platform	• Apple App Store, Huawei AppGallery, Tencent My App
Developer	• Unknown, commercial, psychological service organization, individual developer
Target user	• Children or adolescents, general public, specific for women and patients with mental health problems
Update time	• Days from the retrieval date to the last update
Star rating	• Star rating (out of 5) left by users in the app store
Downloads	• Number of app downloads in the app store
Privacy protection	• Does the description of the app claim to provide privacy protection? (Yes or no) • Is there an obvious privacy protection logo in the process of using apps? (Yes or no) • Does the app report relevant privacy protection regulations? (Yes or no)
Evidence-based and professional background	• Does the app claim to be designed based on proven psychotherapy theory or opinions of mental health service professionals (such as clinicians and psychotherapists) or whether the usability of the app has been proved by peer-reviewed academic research? (Yes or no)

Quality Appraisal of Apps

To evaluated the quality of the apps, we used the Mobile App Rating Scale (MARS), a validated scoring tool for assessing the quality of mHealth apps [23]. MARS has been used to evaluate the quality of different apps, such as apps for mental disorders [24,25], nutrition [26], drug-drug interaction checks [27,28], and chronic disease management [29-31]. MARS contains 23 items, including 4 objective quality subscales of engagement, functionality, aesthetics, and information quality and 1 subjective quality subscale. All items were rated on a 5-point Likert scale from 1 (inadequate) to 5 (excellent). We emphasized the objective quality of the apps, so the subjective quality subscale was excluded from the study. Before formal scoring, all reviewers evaluated the apps that provided only a single mental health service (excluded from the analysis) and discussed inconsistencies in and doubts regarding the results to ensure a unified understanding of MARS projects and standards. To fully experience the service provided by the apps, 2 independent reviewers downloaded and used each app for at least 15 min. The score of each subscale is calculated as the mean of the items in that subscale, and the total score is the mean of each subscale, which describes the overall quality of the app.

Statistical Analysis

The quantitative variables of the MARS score in quality evaluation are described by the mean and SD. Classification variables, such as app characteristics, are described by frequency and percentage. To ensure the reliability of the quality assessment of 2 independent observers, the intragroup correlation coefficients (2-way random, mean measurement, and absolute consistency) were used to evaluate the consistency of commentators at the subscale and overall score level [32]. Pearson correlation coefficients were used to compare (1) the MARS total score and each subscale score, (2) the MARS total score and the user rating, and (3) the user rating and each subscale score. All statistical analyses were conducted using SPSS Statistics 25.

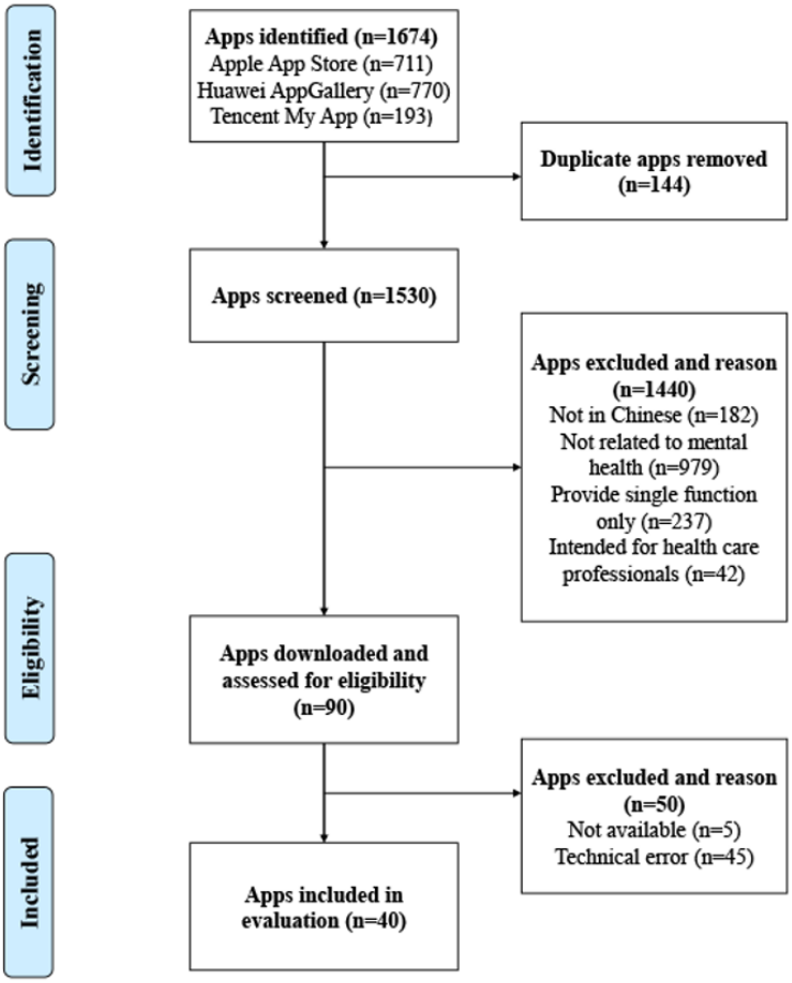
Results

App Selection

A total of 1674 apps were identified through keyword retrieval (711 [42.47%] apps from Apple App Store, 770 [46.00%] apps from Huawei AppGallery, and 193 [11.53%] apps from Tencent My App). Combining the search results of the 3 app stores, 144 (8.6%) duplicate apps were excluded. A total of 1440 (86.02%) apps were excluded on the basis of the exclusion-inclusion criteria. The remaining 90 (5.38%) apps were downloaded onto the evaluation device for further evaluation. Of these 90 apps, 45 (50%) were excluded because of technical reasons (unable to download or use normally because of major technical reasons) and 5 (5.6%) were excluded because of the need to provide special authentication (employees/students). Finally, 40 (44.4%) apps were included in this study (Figure 1).

XSL•FO
RenderX

Figure 1. Flowchart for the systematic search and selection of apps.



General Characteristics

Of the 40 apps included, 31 (77.5%) are from Huawei AppGallery, 6 (15%) from Apple App Store, and 3 (7.5%) from Tencent My App. Most of the apps (35/40, 87.5%) are developed by companies (4 of them are companies mainly engaged in mental health services), 2 (5%) apps are from professional psychological counseling centers, and 3 (7.5%) apps are from individual developers. In addition, 2 of the 40 (5%) apps are specifically designed for adolescents or children. Furthermore, 21 (52.5%) were updated more than 1 month and less than 1 year ago, 15 (37.5%) were maintained within 1 month, and 4 (10%) were updated more than 1 year ago (Table 2).

XSL•FO
RenderX

Table 2. Flowchart for the systematic search and selection of apps (N=40).

Assessment measure	n (%)
Platform	
Apple App Store	6 (15)
Huawei AppGallery	31 (77.5)
Tencent My App	3 (7.5)
Developer	
Commercial	35 (87.5)
Psychological service organization	2 (5)
Individual developer	3 (7.5)
Target user	
Children or adolescents	2 (5)
General public	38 (95)
Specific for women	0
Patients with mental health problems	0
Update time	
>1 year	4 (10)
>1 month and <1 year	21 (52.5)
<1 month	15 (37.5)
Privacy protection	
Privacy protection mentioned in the description	17 (42.5)
An obvious privacy protection logo	23 (57.5)
Relevant privacy protection regulations	33 (82.5)
Evidence-based and professional background[a]	
Proven psychotherapy theory	5 (12.5)
The opinions of mental health service professionals	14 (35)
Peer-reviewed academic research	1 (2.5)
Not mentioned	21 (52.5)

[a]Some apps claim to be designed based on one or more scientific foundations.

Privacy Protection

Of the 40 apps, 17 (42.5%) mentioned the protection of user privacy in the descriptive content of the app store. More than half of the apps (23/40, 57.5%) provide obvious identification during app use to remind users of privacy protection. Almost all apps (33/40, 82.5%) provide links in the app market interface or within the app to access the privacy policy. However, less than two-thirds of these 33 apps (n=21, 63.6%) display privacy policies before users log in to their accounts. All privacy policies explain how to collect user data, and almost all privacy policies (30/33, 90.9%) express how to share, transfer, and publicly disclose user data. Less than half of the apps explain how to store information (16/33, 48.5%) and how to use cookies and other similar technologies (15/33, 45.5%). Of the 33 privacy policies, 14 (42.4%) list the rights of users to manage personal data. For example, users have the right to delete or correct their data. Slightly more than half of the policies (19/33, 57.6%) explain the processing method of minor personal information. Finally, 18 of the 33 (54.5%) apps provide the contact information of the data protection officer that users can access when they have questions or opinions about the content of the privacy policy.

Evidence-Based and Professional Background

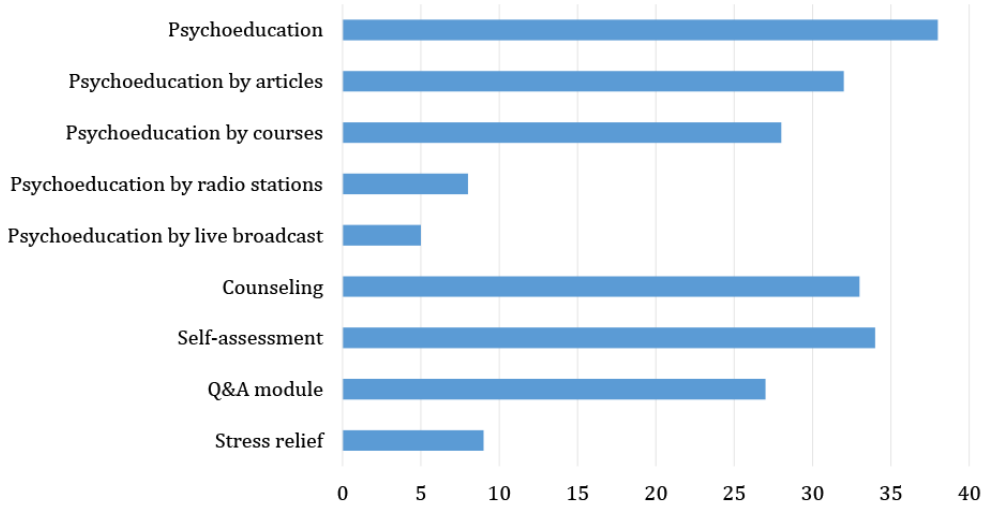
In the description of the app market, 21 of the 40 (52.5%) apps do not mention that the app design specifies a relevant scientific basis. The remaining 19 (47.5%) apps claim to be designed on the basis of 1 or more scientific foundations: 14 (73.7%) apps are described as being designed according to the opinions of mental health professionals, such as clinicians, psychologists, and psychotherapists, while 5 (26.3%) apps claim to be designed using proven psychotherapy theories, such as cognitive behavioral therapy; of these 5 apps, only 1 (20%) claims that its usability has also been confirmed by peer-reviewed academic research.

XSL•FO
RenderX

Functionality Review

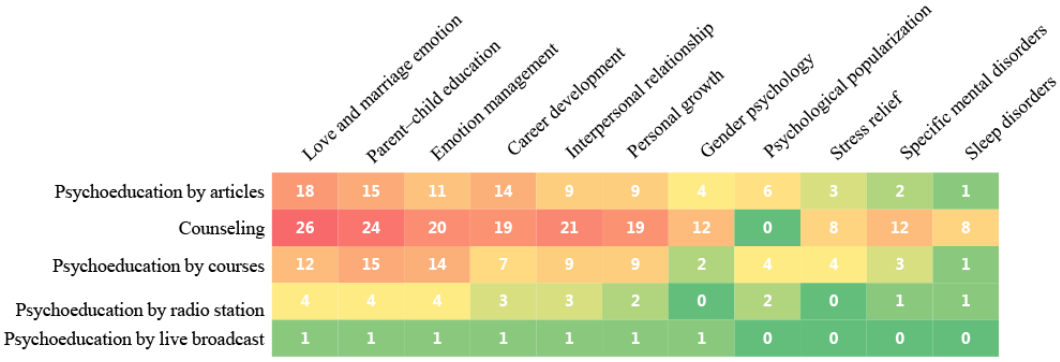
Psychoeducation (38/40, 95%), counseling (33/40, 82.5%), and self-assessment (34/40, 85%) occur in more than three-quarters of the apps, the Q&A community (27/40, 67.5%) appears in about two-thirds of the apps, and stress relief modules (9/40, 22.5%) exist in less than a quarter of the apps (Figure 2). The most common combination in multipurpose apps is the combination of psychoeducation, counseling, self-assessment, and the Q&A community (14/40, 35%). The second is the combination of psychoeducation, counseling, and self-assessment (7/40, 17.5%).

Figure 2. Functional review results of the multipurpose mental health apps. Q&A: question and answer.



Most of the applied psychological education, psychological counseling, and self-assessment function modules set up a topic classification in navigation. To understand the distribution of topics provided in the functional modules of the current apps, we created a heatmap, as shown in Figure 3. The most common themes are love and marriage emotion, parent-child education, and emotion management. The next most common themes are mental disorders in career development, interpersonal relationships, and personal growth.

Figure 3. A heatmap of the topic type (top) for each function and the app function type. Note: numbers in white refer to the frequency of topic types involved in app functions. Warmer colors indicate higher counts.



Psychoeducation

Almost all apps (38/40, 95%) provide psychological education intervention. The mental health education part of the apps is reflected by reading articles related to mental health (32/38, 84.2%), learning relevant courses (28/38, 73.7%), and obtaining relevant information through radio stations (8/38, 21.1%) or live broadcast (5/38, 13.2%).

Among them, the most common way is reading articles related to mental health. Most of these articles are originally created by the platform or psychological counselors, which have guiding and educational significance for the public. More than half of the apps (19/32, 59.4%) have classified the topics of articles to set columns. The main content of the columns includes love and marriage emotion (18/19, 94.7%), parent-child education (15/19, 78.9%), career development (14/19, 73.7%), and emotion management (11/19, 57.9%). Additionally, it also includes themes, such as interpersonal relationships (9/19, 47.4%), personal growth (9/19, 47.4%), gender psychology (4/19, 21.1%), and stress relief (6/19, 31.6%). However, only 2 of the 32 (6.3%) apps have set up columns for specific disorders, such as depression and anxiety, and 1 (3.1%) app has a column for sleep disorders. In addition, 2 (6.3%) apps have columns for students or teenagers, and 1 (3.1%) app set up a column with the rehabilitation story of psychological disorders as the main content.

Mental health–related courses record videos in the form of online education and invite mental health service professionals

XSL•FO
RenderX

to present their professional knowledge, which covers common mental health knowledge, such as emotion management, love emotion, and psychological knowledge. Some of the app courses are free, while most charge a specific fee, ranging from as low as RMB 1.9 (US $0.3) to more than RMB 10,000 (US $1596). Only 6 of 28 (21.4%) apps offer completely free courses. Most other apps (22/28, 78.6%) provide free and paid courses at the same time, and users can choose based on their situation.

Radio and live broadcast are considered less psychological education ways. Mental health education provided using the radio station is mainly manifested in showing users past cases of mental health disorder adjustment or sharing common methods of mental health disorder adjustment. Live broadcast makes up for the shortcomings of the radio form. Users can directly contact consultants through a live broadcast, which strengthens the interaction between users and consultants.

Counseling

Mental health counseling services are provided by mental health service providers, who help solve psychological problems and "heal the soul" through online listening and answering of questions. This functionality is provided in 33 of 40 (82.5%) of the apps. Particularly, the Xiaoxin Psychology app applies artificial intelligence technology to mental health services and provides online counseling services through intelligent robots. Additionally, almost all apps provide consulting services by psychological counselors, and users can select an appropriate provider by viewing the basic information about the psychological counseling provider or modifying the label. Common labels include professional qualification (31/33, 93.9%), areas of expertise (32/33, 97.0%), user evaluation (24/33, 72.7%), and service person-times (23/33, 69.7%). The professional qualification of psychological counselors is the key factor for users to choose from. However, only 19 of 33 (57.6%) apps clearly express the authenticity of professional qualification; 4 of these 19 (21.1%) apps provide evidence of professional psychological counselor qualification, such as a certificate photo or a certificate number. In addition, 15 of the 19 (78.9%) apps are guaranteed by the platform to ensure the authenticity of counselor data. All app downloads are free, but users are charged a specific consulting fee. The consultation cost varies depending on the time or number of times. Almost all consultants (30/33, 90.9%) use voice chat to communicate with consumers. Others provide consultation using text and pictures (21/33, 63.6), video communication (20/33, 60.6%), and offline face-to-face consultation (16/33, 48.5%).

Self-Assessment

The most common function in the apps is psychological testing, accounting for 34 of 40 (85%) of the total. The apps provide some evidence-based or entertainment scales, and users can understand their mental health status through self-assessment of scale problems. Most apps (26/34, 76.5%) provide evidence-based scales, the most common of which include the Self-Rating Anxiety Scale (24/26, 92.3%) [33], the Self-Rating Depression Scale (22/26, 84.6%) [34], and the Symptom Checklist-90-Revised (14/26, 53.8%) [35]. Almost all apps (32/34, 94.1%) provide a scale for the nature of entertainment to attract users' attention. These scales are developed by the app team or formed by a scale with unclear origin to evaluate the user's emotion, personality, ability, sleep status, professional interest, and interpersonal status.

Q&A Module

The Q&A community embodies the great advantages of mental health apps compared with traditional mental health services. Of the 40 apps, 27 (67.5%) provide this module. The Q&A community gathers other users and consultants on the same platform for rapid communication of mental health problems between users and between users and consultants, which is difficult to achieve by traditional psychological services. Users express their troubles, puzzles, or problems in the Q&A community, discuss and communicate with other users and consultants through Q&A feedback, solve problems, and gain knowledge.

Stress Relief

The stress relief module regulates the user's mood, improves the sleep state, and relieves the user's psychological pressure through proven ways, such as meditation and audio decompression. Of the 40 apps, 9 (22.5%) provide functional modules for stress relief. Meditation is considered a popular intervention method to relieve stress. Of these 9 apps, 7 (77.8%) provide functional modules to assist meditation. The modules guide stress relief training based on mindfulness or breathing technology (7/7, 100%) and cognitive behavioral therapy (1/7, 14.3%) in the form of audio or video. Audio is the main medium to help users relieve pressure. In addition to the audio used for meditation, other types of audio include light music (4/9, 44.4%), autonomous sensor meridian response audio (1/9, 11.1%), and nature recording (4/9, 44.4%). Additionally, 4 (44.4%) apps specifically provide audio to help users sleep.

MARS Evaluation

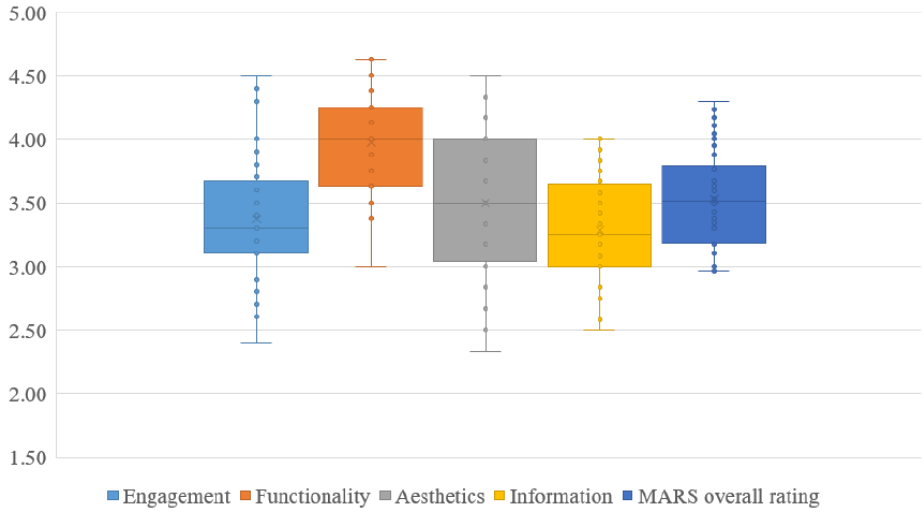
The overall MARS score showed high interreviewer reliability (intraclass correlation coefficient [ICC] 0.95, 95% CI 0.858-0.960). Simultaneously, all subscales also showed good consistency: engagement ICC 0.93 (95% CI 0.865-0.963), functionality ICC 0.71 (95% CI 0.462-0.847), aesthetics ICC 0.91 (95% CI 0.833-0.953), and information ICC 0.85 (95% CI 0.727-0.923).

The total average MARS score of all apps was 3.54 (SD 0.39), and the total score ranged from 2.96 (Enmasa Psychology) to 4.30 (Yi Psychology). The MARS score of 7 of 40 (17.5%) apps was ≥4. Furthermore, 30 of 40 (75%) apps had MARS scores ranging from 3.0 to 3.99. MARS scores of 3 of 40 (7.5%) apps ranged from 2.0 to 2.99. There were no apps with a score of <2.

The average scores of each subscale were as follows: information quality score=3.29 (SD 0.41), engagement quality score=3.37 (SD 0.51), aesthetic quality score=3.50 (SD 0.61), and functional quality score=3.97 (SD 0.37). The aesthetic quality score showed the largest span, with a minimum of 2.33 and a maximum of 4.50. The information quality part was the lowest, ranging from 2.50 to 4.00. The rating distribution of overall quality and 4 subscale dimensions is shown in Figure 4.

XSL•FO
RenderX

Figure 4. Graphical representation of the distribution of the MARS overall and subscale score. The median, the interquartile distance, and the range were given (N = 40). MARS: Mobile App Rating Scale.



The overall MARS score was significantly positively correlated with the scores of each subscale (r=0.62-0.88, *P*<.001). However, the user rating of the app market was not significantly correlated with the total MARS score (r=0.17, *P*=.33) and the scores of various scales (r=0.05-0.22, *P*=.21–.77; Table 3).

Combined with the professional background of app development, the total average MARS score of the app described as designed according to the opinions of mental health professionals is 3.66. The total average score of apps claiming to use proven psychotherapy theory is 3.77. The only app that clearly states that its usability is confirmed by peer-reviewed academic research has a total MARS score of 3.95. The quality score for all of the above cases is higher than the overall average score for all apps. However, the total average score of apps without reference to app design involving relevant scientific basis is 3.40.

Table 3. Correlation between the Mobile App Rating Scale subscale and the overall score and the user star score.

Characteristic	Engagement, correlation (*P* value)	Functionality, correlation (*P* value)	Aesthetics, correlation (*P* value)	Information, correlation (*P* value)	Overall rating, correlation (*P* value)
Engagement	—[a]	—	—	—	—
Functionality	0.39 (.01)	—	—	—	—
Aesthetics	0.67 (<.001)	0.31 (.049)	—	—	—
Information	0.73 (<.001)	0.51 (.001)	0.65 (<.001)	—	—
Overall rating	0.88 (<.001)	0.62 (<.001)	0.86 (<.001)	0.88 (<.001)	—
User star rating[b]	0.22 (.21)	0.21 (.23)	0.08 (.64)	0.05 (.77)	0.17 (.33)

[a]Not applicable.

[b]Apps with zero user star ratings were excluded.

Discussion

Principal Findings

This study identified 40 multipurpose mental health apps, understood their main functional distribution, and evaluated their content and quality.

The evaluation of the professional background of app development provides an opportunity for app developers to improve the scientificity and accuracy of these apps. Among all included apps, only 2 (5%) are developed by professional organizations engaged in mental health services, while most of the apps (38/40, 95%) are developed by commercial companies or individuals without mental health–related backgrounds. Additionally, more than half of the app development background lacks the participation of professionals or scientific theories. In the description of the apps, only 1 (2.5%) clearly stated that its availability had been confirmed by peer-reviewed academic research. The absence of the development process in the app description may raise questions about the credibility of the apps [36]. Simultaneously, the lack of a professional development background may lead to an inappropriate final app, which is considered a potential threat to users. Previous evidence also emphasizes that the development process should include the participation of health care professionals and target users with regard to apps that provide health content and collect health data [37]. Therefore, these findings emphasize the need to take action to ensure the scientific quality of mental health apps, which will improve the reliability and quality of the content provided by the apps.

XSL•FO
RenderX

The in-depth analysis revealed that the most common functional combination of multipurpose mental health apps is psychological education, counseling, self-assessment, and the Q&A community. Only 1 (2.5%) app contains all functional modules, and its quality score is also good. In all apps, psychological education, self-assessment, and counseling occupy the main positions. However, the online consultation function of most apps is provided by professional mental health care personnel. Therefore, the problem of insufficient mental health service personnel still exists. However, we found in this search an app, Xiaoxin Psychology, that provides the possibility of solving this problem. Xiaoxin Psychology combines artificial intelligence with mental health to replace mental health service personnel. Although we have not found a test of the effectiveness of the app, studies have confirmed the effectiveness of evidence-based computerized interventions in alleviating anxiety and depression in adults [38].

Among the 40 apps identified, except "grape heart," which is an app for children with autism, no app for a specific mental disorder was found. This is different from foreign mental health apps [39-41]. However, we found that the included apps classify the service theme and set up navigation in the menu for quick access. We observed that the obstacles of love and marriage emotion and parent-child education are the most common. This may be related to the traditional Chinese concept of paying attention to family emotion. COVID-19 has increased the demand for mental health services, but only 3 (7.5%) of all assessment apps have added psychological aid plates. This may be due to an untimely update of the current apps. App developers set content classification modules on the basis of current events and hot spots, which is a good way to attract new users and stabilize old users.

The 40 apps' choice of target population also has specific characteristics. Only 2 (5%) of the apps are designed specifically for teenagers or children, and there are no apps designed for women. App developers prefer ordinary adult users, which may be because it is easier to obtain users and maintain the stability of users. However, adolescents are one of the most vulnerable to mental health problems [42]. They are often reluctant to seek professional help because of their sense of shame and tendency toward self-reliance [8]. The method of getting help based on a network provides a way to overcome these obstacles [43-45].

Although almost all apps provide privacy policies for the collection and use of users' personal data, they lack detailed information about data storage, user management permissions, and the use of cookies. Moreover, in this evaluation, most apps lacked a description of the endpoint of data sharing, which is consistent with the results of previous studies [46]. The economic benefits of shared data promote the occurrence of such situations and pose a threat of user data disclosure [47]. However, personal health information is highly sensitive, and the disclosure of health information may cause varying degrees of negative effects and even death [48]. Additionally, people may refuse to use mHealth apps because of concerns about health data security and privacy [49,50]. Furthermore, concerns about privacy protection are exacerbated by people's sense of shame about using mental health services [51-53]. However, the trust between app developers and users may be damaged by the lack of clarity of privacy policies, which results in the loss of potential long-term users [54]. Therefore, there is still a long way to go in terms of the compliance with privacy policy content and the pertinence to special types of apps (such as health).

The overall quality of the multipurpose mental health apps we reviewed is good. The MARS score ranges from 2.96 to 4.30. There is a gap in the quality of the apps, which is similar to that in previous studies [55]. However, the 3 (7.5%) apps with the highest MARS quality score also have the highest download frequency, indicating the attraction of high-quality apps to target users. Additionally, the apps show advantages in functional evaluation rather than the information part. This emphasizes that future app development should focus on improving the information quality of apps. Adding professional mental health care personnel to the app development process may be a feasible way. This will also provide a reference for users to evaluate the degree of expertise involved in the app development process before downloading [56]. In this study, apps with professional development backgrounds also reflected their advantages in quality scoring. The user star rating was not correlated with the total MARS quality score, which may indicate that there is a different structure between the user star rating and the app quality score. Similarly, in previous studies, it has been reported that app quality depends not only on the content but also on the function and design method of the content [57]. Furthermore, the star rating of the app market may involve the early version of the apps and cannot fully represent the current version, which may lead to distortion in the evaluation of the current version [58].

Contribution

This study conducted a specific survey on the content and quality of multipurpose mental health apps in China based on a systematic and evidence-based approach. To the best of our knowledge, this is the first attempt to incorporate the professional background of app development and the user privacy protection policy into the evaluation of mental health apps in China. These findings will assist app developers in enhancing current apps or design new apps. Additionally, through the analysis of privacy policies, users can better understand the potential risks of providing information to service providers. Furthermore, we found that the combination of artificial intelligence and mental health may provide the possibility of solving the problem of insufficient mental health service personnel. This will be the direction of designing and creating apps in the future.

Limitations and Future Work

There are some limitations to this study. We may have missed some apps. Keywords retrieval cannot exhaust all apps. Some apps that met the inclusion criteria were ignored because the title or description did not contain search criteria related to mental health. Moreover, the app market is constantly changing, new apps may be on the shelf at any time, and old apps may be deleted for various reasons. Although the research examines the apps' emphasis on user privacy protection, we cannot verify whether the apps really implement privacy protection measures. Additionally, this study did not verify the scientificity of the content provided by the apps, and in-depth research will be

XSL•FO
RenderX

continued in the future. Two researchers independently screened the eligibility of the apps, extracted the characteristics of the apps, and used MARS to evaluate the quality of the apps. The rater's reliability is good or excellent, but if more researchers participate, the results may be more objective. Hence, we will consider increasing the number of researchers in follow-up studies. Furthermore, some of our findings may reveal the direction of such research in the future. We found some behavior change techniques aimed at improving users' mental health, such as meditation. Future research can further evaluate the quality and characteristics of behavior change techniques in these apps. In addition, this survey also identified some high-quality apps. Before widely recommending these apps, a further randomized controlled trial can be used to determine and compare their effectiveness.

Conclusion

This study identified 40 multipurpose mental health apps, analyzed their main functional distribution, and evaluated their content and quality. These findings will assist app developers in enhancing current apps or design new apps. The quality of multipurpose mental health apps in China's main app markets is generally good. Most apps provide rich functionality and classify the service theme to set up navigation in the menu for quick access. However, the lack of professional background in the app development process raises concerns about the scientificity of the apps. Furthermore, the privacy protection policy of the apps also needs to be described in more detail.

Conflicts of Interest

None declared.

References

1. World Health Organization. Global Challenge for Movement on Mental Health Kicks Off as Lack of Investment in Mental Health Leaves Millions without Access to Services. URL: https://www.who.int/news/item/07-10-2020-global-challenge-for-movement-on-mental-health-kicks-off-as-lack-of-investment-in-mental-health-leaves-millions-without-access-to-services [accessed 2020-08-12]
2. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yan J, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. Lancet Psychiatry 2019 Mar;6(3):211-224. [doi: 10.1016/S2215-0366(18)30511-X] [Medline: 30792114]
3. Yin H, Wardenaar KJ, Xu G, Tian H, Schoevers RA. Help-seeking behaviors among Chinese people with mental disorders: a cross-sectional study. BMC Psychiatry 2019 Nov 29;19(1):373 [FREE Full text] [doi: 10.1186/s12888-019-2316-z] [Medline: 31783825]
4. Zhao K, He Y, Zeng Q, Ye L. Factors of mental health service utilization by community-dwelling adults in Shanghai, China. Community Ment Health J 2019 Jan;55(1):161-167. [doi: 10.1007/s10597-018-0352-7] [Medline: 30488354]
5. Xia L, Jiang F, Rakofsky J, Zhang Y, Shi Y, Zhang K, et al. Resources and workforce in top-tier psychiatric hospitals in China: a nationwide survey. Front Psychiatry 2021;12:573333 [FREE Full text] [doi: 10.3389/fpsyt.2021.573333] [Medline: 33716804]
6. Sun KS, Lam TP, Wu D. Chinese perspectives on primary care for common mental disorders: barriers and policy implications. Int J Soc Psychiatry 2018 Aug;64(5):417-426. [doi: 10.1177/0020764018776347] [Medline: 29781372]
7. Shi W, Shen Z, Wang S, Hall BJ. Barriers to professional mental health help-seeking among Chinese adults: a systematic review. Front Psychiatry 2020;11:442 [FREE Full text] [doi: 10.3389/fpsyt.2020.00442] [Medline: 32508688]
8. Salaheddin K, Mason B. Identifying barriers to mental health help-seeking among young adults in the UK: a cross-sectional survey. Br J Gen Pract 2016 Oct;66(651):e686-e692 [FREE Full text] [doi: 10.3399/bjgp16X687313] [Medline: 27688518]
9. Srivastava K, Chatterjee K, Bhat PS. Mental health awareness: the Indian scenario. Ind Psychiatry J 2016;25(2):131-134 [FREE Full text] [doi: 10.4103/ipj.ipj_45_17] [Medline: 28659690]
10. Niendam TA, Tully LM, Iosif A, Kumar D, Nye KE, Denton JC, et al. Enhancing early psychosis treatment using smartphone technology: a longitudinal feasibility and validity study. J Psychiatr Res 2018 Jan;96:239-246. [doi: 10.1016/j.jpsychires.2017.10.017] [Medline: 29126059]
11. Xiang Y, Yang Y, Li W, Zhang L, Zhang Q, Cheung T, et al. Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. Lancet Psychiatry 2020 Mar;7(3):228-229 [FREE Full text] [doi: 10.1016/S2215-0366(20)30046-8] [Medline: 32032543]
12. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
13. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. Int J Soc Psychiatry 2020 Jun;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
14. Pérez Sust P, Solans O, Fajardo JC, Medina Peralta M, Rodenas P, Gabaldà J, et al. Turning the crisis into an opportunity: digital health strategies deployed during the COVID-19 outbreak. JMIR Public Health Surveill 2020 May 04;6(2):e19106 [FREE Full text] [doi: 10.2196/19106] [Medline: 32339998]

XSL•FO
RenderX

15. Kola L. Global mental health and COVID-19. Lancet Psychiatry 2020 Aug;7(8):655-657 [FREE Full text] [doi: 10.1016/S2215-0366(20)30235-2] [Medline: 32502468]
16. Tan Y, Teng Z, Qiu Y, Tang H, Xiang H, Chen J. Potential of mobile technology to relieve the urgent mental health needs in China: web-based survey. JMIR Mhealth Uhealth 2020 Jul 07;8(7):e16215 [FREE Full text] [doi: 10.2196/16215] [Medline: 32673239]
17. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
18. Yin H, Wardenaar KJ, Wang Y, Wang N, Chen W, Zhang Y, et al. Mobile mental health apps in China: systematic app store search. J Med Internet Res 2020 Jul 27;22(7):e14915 [FREE Full text] [doi: 10.2196/14915] [Medline: 32716301]
19. Shang J, Wei S, Jin J, Zhang P. Mental health apps in China: analysis and quality assessment. JMIR Mhealth Uhealth 2019 Nov 07;7(11):e13236 [FREE Full text] [doi: 10.2196/13236] [Medline: 31697245]
20. Balebako R, Cranor L. Improving app privacy: nudging app developers to protect user privacy. IEEE Secur Privacy 2014 Jul;12(4):55-58. [doi: 10.1109/msp.2014.70]
21. AppInChina. The AppInChina App Store Index. URL: https://www.appinchina.co/market/app-stores/ [accessed 2020-12-10]
22. Canalys. China Smartphone Market Q4 2020 and Full Year 2020. URL: https://www.canalys.com/newsroom/china-smartphone-market-declines-11-in-2020 [accessed 2021-08-15]
23. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
24. Lau N, O'Daffer A, Yi-Frazier JP, Rosenberg AR. Popular evidence-based commercial mental health apps: analysis of engagement, functionality, aesthetics, and information quality. JMIR Mhealth Uhealth 2021 Jul 14;9(7):e29689 [FREE Full text] [doi: 10.2196/29689] [Medline: 34259639]
25. Myers A, Chesebrough L, Hu R, Turchioe MR, Pathak J, Creber RM. Evaluating commercially available mobile apps for depression self-management. AMIA Annu Symp Proc 2020;2020:906-914 [FREE Full text] [Medline: 33936466]
26. Mandracchia F, Llauradó E, Tarro L, Valls RM, Solà R. Mobile phone apps for food allergies or intolerances in app stores: systematic search and quality assessment using the Mobile App Rating Scale (MARS). JMIR Mhealth Uhealth 2020 Sep 16;8(9):e18339 [FREE Full text] [doi: 10.2196/18339] [Medline: 32936078]
27. Shen C, Jiang B, Yang Q, Wang C, Lu KZ, Gu M, et al. Mobile apps for drug-drug interaction checks in Chinese app stores: systematic review and content analysis. JMIR Mhealth Uhealth 2021 Jun 15;9(6):e26262 [FREE Full text] [doi: 10.2196/26262] [Medline: 33962910]
28. Kim BY, Sharafoddini A, Tran N, Wen EY, Lee J. Consumer mobile apps for potential drug-drug interaction check: systematic review and content analysis using the Mobile App Rating Scale (MARS). JMIR Mhealth Uhealth 2018 Mar 28;6(3):e74 [FREE Full text] [doi: 10.2196/mhealth.8613] [Medline: 29592848]
29. Kalhori SRN, Hemmat M, Noori T, Heydarian S, Katigari MR. Quality evaluation of English mobile applications for gestational diabetes: app review using Mobile Application Rating Scale (MARS). Curr Diabetes Rev 2021;17(2):161-168. [doi: 10.2174/1573399816666200703181438] [Medline: 32619173]
30. Gong E, Zhang Z, Jin X, Liu Y, Zhong L, Wu Y, et al. Quality, functionality, and features of Chinese mobile apps for diabetes self-management: systematic search and evaluation of mobile apps. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e14836 [FREE Full text] [doi: 10.2196/14836] [Medline: 32255432]
31. Knitza J, Tascilar K, Messner E, Meyer M, Vossen D, Pulla A, et al. German mobile apps in rheumatology: review and analysis using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2019 Aug 05;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
33. Jegede RO. Psychometric attributes of the Self-Rating Anxiety Scale. Psychol Rep 1977 Feb;40(1):303-306. [doi: 10.2466/pr0.1977.40.1.303] [Medline: 840986]
34. Zung WW. A self-rating depression scale. Arch Gen Psychiatry 1965 Jan;12:63-70. [doi: 10.1001/archpsyc.1965.01720310065008] [Medline: 14221692]
35. Derogatis L. SCL-90-R: Administration, scoring, and procedures manual I for the R(evised) version. Clin Psychometr Res 1977:A.
36. Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E. Mobile health apps for self-management of rheumatic and musculoskeletal diseases: systematic literature review. JMIR Mhealth Uhealth 2019 Nov 26;7(11):e14730 [FREE Full text] [doi: 10.2196/14730] [Medline: 31769758]
37. Najm A, Nikiphorou E, Kostine M, Richez C, Pauling JD, Finckh A, et al. EULAR points to consider for the development, evaluation and implementation of mobile health applications aiding self-management in people living with rheumatic and musculoskeletal diseases. RMD Open 2019;5(2):e001014 [FREE Full text] [doi: 10.1136/rmdopen-2019-001014] [Medline: 31565245]

XSL•FO
RenderX

38. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
39. Marshall JM, Dunstan DA, Bartik W. Apps with maps-anxiety and depression mobile apps with evidence-based frameworks: systematic search of major app stores. JMIR Ment Health 2020 Jun 24;7(6):e16525 [FREE Full text] [doi: 10.2196/16525] [Medline: 32579127]
40. Qu C, Sas C, Daudén Roquet C, Doherty G. Functionality of top-rated mobile apps for depression: systematic search and evaluation. JMIR Ment Health 2020 Jan 24;7(1):e15321 [FREE Full text] [doi: 10.2196/15321] [Medline: 32012079]
41. Martinengo L, Van Galen L, Lum E, Kowalski M, Subramaniam M, Car J. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. BMC Med 2019 Dec 19;17(1):231 [FREE Full text] [doi: 10.1186/s12916-019-1461-z] [Medline: 31852455]
42. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry 2015 Mar;56(3):345-365. [doi: 10.1111/jcpp.12381] [Medline: 25649325]
43. Ridout B, Campbell A. The use of social networking sites in mental health interventions for young people: systematic review. J Med Internet Res 2018 Dec 18;20(12):e12244 [FREE Full text] [doi: 10.2196/12244] [Medline: 30563811]
44. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: systematic narrative review. J Med Internet Res 2019 Nov 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
45. Stunden C, Zasada J, VanHeerwaarden N, Hollenberg E, Abi-Jaoudé A, Chaim G, et al. Help-seeking behaviors of transition-aged youth for mental health concerns: qualitative study. J Med Internet Res 2020 Oct 05;22(10):e18514 [FREE Full text] [doi: 10.2196/18514] [Medline: 33016882]
46. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. JAMA Netw Open 2019 Apr 05;2(4):e192542 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2542] [Medline: 31002321]
47. Martinengo L, Stona A, Griva K, Dazzan P, Pariante CM, von Wangenheim F, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. J Med Internet Res 2021 Jul 30;23(7):e27619 [FREE Full text] [doi: 10.2196/27619] [Medline: 34328431]
48. Rosa BMG, Anastasova S, Yang GZ. NFC-powered implantable device for on-body parameters monitoring with secure data exchange link to a medical blockchain type of network. IEEE Trans Cybern 2021 Jul 01. [doi: 10.1109/TCYB.2021.3088711] [Medline: 34197334]
49. Kao C, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
50. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and security in mobile health: a research agenda. Computer (Long Beach Calif) 2016 Jun;49(6):22-30 [FREE Full text] [doi: 10.1109/MC.2016.185] [Medline: 28344359]
51. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
52. Zhou L, Parmanto B. User preferences for privacy protection methods in mobile health apps: a mixed-methods study. Int J Telerehabil 2020 Dec 08;12(2):13-26 [FREE Full text] [doi: 10.5195/ijt.2020.6319] [Medline: 33520091]
53. Kotz D. A threat taxonomy for mHealth privacy. 2011 Presented at: International Conference on Communication Systems & Networks; 2011; Bangalore.
54. Rosenfeld L, Torous J, Vahia IV. Data security and privacy in apps for dementia: an analysis of existing privacy policies. Am J Geriatr Psychiatry 2017 Aug;25(8):873-877. [doi: 10.1016/j.jagp.2017.04.009] [Medline: 28645535]
55. Song Y, Chen H. Evaluating Chinese mobile health apps for ankylosing spondylitis management: systematic app search. JMIR Mhealth Uhealth 2021 Jul 14;9(7):e27234 [FREE Full text] [doi: 10.2196/27234] [Medline: 34259644]
56. Visvanathan A, Hamilton A, Brady RRW. Smartphone apps in microbiology: is better regulation required? Clin Microbiol Infect 2012 Jul;18(7):E218-E220 [FREE Full text] [doi: 10.1111/j.1469-0691.2012.03892.x] [Medline: 22563840]
57. Wilks CR, Chu C, Sim D, Lovell J, Gutierrez P, Joiner T, et al. User engagement and usability of suicide prevention apps: systematic search in app stores and content analysis. JMIR Form Res 2021 Jul 14;5(7):e27018 [FREE Full text] [doi: 10.2196/27018] [Medline: 34259163]
58. Mojica Ruiz I, Nagappan M, Adams B, Berger T, Dienst S, Hassan A. An examination of the current rating system used in mobile app stores. IEEE Softw 2015;55(1):1-1. [doi: 10.1109/MS.2017.265094809]

Abbreviations

ICC: intraclass correlation coefficient
MARS: Mobile Application Rating Scale
mHealth: mobile health
Q&A: question and answer

XSL•FO
RenderX

Edited by G Eysenbach; submitted 05.10.21; peer-reviewed by C Pretorius; comments to author 27.10.21; accepted 30.11.21; published 04.01.22.

Please cite as:
Wu X, Xu L, Li P, Tang T, Huang C
Multipurpose Mobile Apps for Mental Health in Chinese App Stores: Content Analysis and Quality Evaluation
JMIR Mhealth Uhealth 2022;10(1):e34054
URL: https://mhealth.jmir.org/2022/1/e34054
doi:10.2196/34054
PMID:34982717

©Xiaoqian Wu, Lin Xu, PengFei Li, TingTing Tang, Cheng Huang. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 04.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Russian-Language Mobile Apps for Reducing Alcohol Use: Systematic Search and Evaluation

Anna Bunova[1]; Veronika Wiemker[2]; Boris Gornyi[1], PhD; Carina Ferreira-Borges[3], PhD; Maria Neufeld[3,4,5], MSc

[1]National Research Center for Therapy and Preventive Medicine of the Ministry of Health of the Russian Federation, Moscow, Russian Federation
[2]Heidelberg Institute of Global Health, Medical Faculty and University Hospital, Heidelberg University, Heidelberg, Germany
[3]WHO European Office for the Prevention and Control of Noncommunicable Diseases, Moscow, Russian Federation
[4]Institute for Mental Health Policy Research, Centre for Addiction and Mental Health, Toronto, ON, Canada
[5]Institute of Clinical Psychology and Psychotherapy, Dresden University of Technology, Dresden, Germany

Corresponding Author:
Anna Bunova
National Research Center for Therapy and Preventive Medicine of the Ministry of Health of the Russian Federation
Petroverigskiy Pereulok 10
Moscow, 101990
Russian Federation
Phone: 7 9151416154
Email: asbunova@gmail.com

Abstract

Background: Personalized prevention tools such as mobile apps designed to reduce alcohol consumption are widespread in mobile app stores accessible in Russia. However, the quality and content of these mobile apps have not been systematically evaluated.

Objective: This study aimed to identify Russian-language mobile apps for reducing alcohol use and to evaluate their quality and potential to change alcohol-related health behavior. It further aimed to identify apps that could facilitate screening and brief interventions in primary health care in Russia.

Methods: A systematic search for mobile apps available in Russia was carried out between April 1 and 15, 2020, December 1 and 15, 2020, and in March 2021 in the iPhone App Store, Google Play Store, and the 4PDA forum. App quality was assessed using the Mobile App Rating Scale (MARS), and structured searches in electronic libraries and bibliographic databases were used to evaluate the apps' evidence base. The number of features facilitating changes in lifestyle behavior was assessed using the App Behavior Change Scale (ABACUS).

Results: We identified 63 mobile apps for reducing alcohol use. The mean MARS quality ratings were high for the subscales of functionality (3.92 out of 5, SD 0.58) and aesthetics (2.96, SD 0.76) and low for engagement (2.42, SD 0.76) and information (1.65, SD 0.60). Additional searches in electronic libraries and bibliographic databases (eLibrary, CyberLeninka, Google Scholar) yielded no studies involving the identified apps. ABACUS scores ranged from 1 to 15 out of 25, with a mean of 5 (SD 3.24). Two of the identified apps might be useful for screening and brief interventions in Russian primary health care after improvements in content and scientific testing.

Conclusions: Russian-language mobile apps for reducing alcohol use are accessible in the app stores. Many of them are aesthetically pleasing, functional, and easy to use. However, information about their scientific trialing or testing is lacking. Most apps contain a low number of features that facilitate changes in lifestyle behavior. Further research should examine the context of Russian-language mobile apps for reducing alcohol use. Our findings underline the need to develop evidence-based apps to mitigate alcohol consumption in Russia and elsewhere.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020167458; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=167458

(JMIR Mhealth Uhealth 2022;10(1):e31058) doi:10.2196/31058

KEYWORDS

alcohol; mHealth; mobile applications; screening and brief intervention; Mobile Application Rating Scale; App Behavior Change Scale; mobile phone

XSL•FO
RenderX

Introduction

Background

Alcohol is one of the leading risk factors contributing to the global burden of disease and mortality [1-3]. The World Health Organization (WHO) estimates that globally, about 3 million deaths are caused by alcohol use each year, almost 1 million of which occur in the WHO European Region as the region with the highest level of per capita alcohol consumption. Drinking alcohol contributes to the development of more than 200 diseases and injuries. It increases the risk of cardiovascular and digestive diseases, neoplasms, mental and behavioral disorders (not limited to alcohol use disorders) as well as violent crimes, suicides, and road traffic accidents [4,5]. The impact of alcohol on mortality in Russia has been well documented, and Russia remains one of the countries with the largest alcohol-attributable burden of diseases worldwide, although substantial improvements were made over the last decade [6-9]. The WHO has recently launched the SAFER initiative to reduce alcohol-related harm, which recommends that health services should provide prevention and treatment interventions to individuals and families at risk of or affected by alcohol use disorders and associated conditions [10]. One of the 5 high-impact interventions of SAFER is the screening and brief intervention (SBI) programs in primary health care (PHC) [10,11].

There is a large body of research supporting the effectiveness of SBI in reducing alcohol consumption and other alcohol-related outcomes [12-15]. Attempts to introduce SBI in the Russian PHC for patients at risk for harmful alcohol use began in 2013 when legislative changes allowed the establishment of SBI as a part of Russia's dispanserization program within PHC facilities [16]. Dispanserization is a set of standardized measures in PHC that includes preventive medical examination for assessing the state of health and is carried out in relation to certain groups of the population in accordance with the legislation of the Russian Federation. Following the currently established provisions, dispanserization includes an evidence-based 2-step screening procedure aiming to provide early and timely detection of conditions and diseases as well as risk factors for their development, including the nonmedical use of drugs and psychotropic substances [17]. The introduced SBI as within the broader dispanserization framework consists of 2 steps [17]. In the first step, the self-administered 3-item version of the Alcohol Use Disorders Identification Test-Consumption (AUDIT-C) is used to detect patients at risk [17,18]. If the screening result is positive, that is, the score exceeds the sex-specific cut-off, the patient is asked to complete the full 10-item version of the Alcohol Use Disorders Identification Test (AUDIT) as part of an interview with a health care professional. The health care professional then provides a brief intervention depending on the results [17,19]. Despite the effectiveness of SBI, there are several barriers to their widespread implementation in PHC settings in Russia beyond the dispanserization framework [20-22]. The AUDIT is based on the concept of a standard drink, which was introduced in several countries worldwide as a measure of alcohol consumption and to provide information about alcohol consumption to consumers, which was also as part of communicating the number of standard drinks on labels of alcoholic beverages [19]. In practice, using the standard drink concept remains a challenge for PHC professionals. PHC professionals report that the concept is not understandable for patients and difficult to calculate with, especially for patients that engage in heavy episodic drinking [23,24]. Moreover, delivering a brief intervention requires specific skills and knowledge as well as additional time and resources from the PHC professionals [20,21].

The development of electronic systems to deliver or support SBIs can potentially address some of these challenges and support health care workers. For instance, electronic devices such as smartphones and tablets can be used instead of the traditional paper-and-pencil screening tests and facilitate counting standard drinks as part of the risk assessment and support the delivery of brief interventions. Moreover, electronic SBIs are also potentially more flexible and can be adapted to take into account the regional patterns of alcohol consumption and make the assessment more personalized. They can also potentially reach larger audiences beyond the health sector [25,26].

The growing popularity of mobile phones and the active development of mobile internet in all regions of Russia open up great opportunities for using mobile apps as tools to change individual health behavior [27]. Mobile apps can provide an additional resource for preventive interventions catering to at-risk populations. However, to be successful, such interventions require the health care professional to select effective, evidence-based, and field-proven mobile apps [28]. A study by Abroms and colleagues [29] showed that making such a choice is difficult since many mobile apps contain inaccuracies and low-quality information, are not tested in practice, or lack an evidence base. Abroms et al [29] point out the potential dangers of such apps, ranging from misinformation to misleading risk level estimates. A rigid evaluation of apps for reducing alcohol use is therefore of great interest to both alcohol consumers and health care professionals. While former studies have described the features of highly rated Russian-language apps for reducing alcohol use, they did not evaluate the app quality and the potential to change alcohol-related health behavior by using validated instruments [30,31]. By closing this gap and providing researchers and health care professionals with an overview of the currently available evidence-based apps for reducing alcohol use, this study may contribute to facilitating the provision of SBI programs in the Russian PHC.

Objective

The aim of this study was to conduct a systematic search and evaluation of Russian-language mobile apps for reducing alcohol use. The specific objectives were to (1) create an overview and establish a list of relevant apps available in Russia, (2) assess their overall quality and evidence base, and (3) evaluate if any of the available apps could be used to support the provision of AUDIT-based SBI in Russia and its broader implementation in PHC facilities.

XSL•FO
RenderX

Methods

Study Design

The study was performed in 2 steps. In step 1, we conducted a systematic app store search to identify the apps for reducing alcohol use, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [32]. In step 2, we evaluated the identified apps by using the Mobile App Rating Scale (MARS) [33] and the App Behavior Change Scale (ABACUS) [34]. For specific steps of the rating procedure, please see below. The study protocol was published in PROSPERO (Prospective Register of Systematic Reviews) [35] under registration CRD42020167458 (review ongoing).

Step 1: Systematic App Store Searches

Eligibility Criteria

We defined mobile apps for reducing alcohol use as tools for tablets or smartphones that facilitate behavioral change related to alcohol use. We excluded apps that were clearly not aimed at the reduction of alcohol use, such as games, barcode scanners as part of the Unified State Automated Information System (EGAIS) tracking alcohol distribution and sales under the Russian Federal Service for Alcohol Market Regulation [36], recipes for alcoholic drinks, and wallpapers. Only currently available and working Russian-language apps with at least a basic or trial version free of charge were included.

Search Strategy

Six systematic searches in the iPhone App Store, Google Play Store, and a Russian internet forum of mobile apps, that is, 4PDA [37] were conducted between April 1 and 15, 2020, December 1 and 15, 2020, and in March 2021. Two native Russian speakers conducted these 6 independent searches on different dates to account for additional mobile apps that were created during the COVID-19 pandemic. Keywords included the Russian words for “alcohol,” “alcoholic drinks,” “spirits,” “beer,” “vodka,” “drink alcohol,” “alcohol calculator,” “alcohol tracker,” “sober,” “alcohol monitoring,” and “breathalyzer” (details in PROSPERO protocol [35] and Multimedia Appendix 1) and were entered through the general search bar of the app stores and the forum.

Screening and Selection of Apps

In the first step, we recorded the name, app icon, developer, store, platform, brief description, and URL of all the available alcohol-related mobile apps. Next, duplicates were removed, and app store descriptions were screened against inclusion criteria. We retained only 1 record if identical versions of an app were available for Android and iPhone operating systems (iOSs). All remaining apps were downloaded onto the study devices (Samsung Galaxy Tab A 7.0 SM-T285 8GB/Android version 9, Lenovo Tablet TB-X704L 64G/Android version 7, and iPhone 11/iOS version 14.0.1). Apps that could not be opened on these devices were excluded.

Data Extraction

The following information was extracted for all the included apps: app name, the app’s star rating on the platform, number of installations, developer, current version, number of ratings for current version, last update, existence of a basic version and paid premium versions, and platform. All included apps were available and all data were updated in the last week of March 2021.

Step 2: Evaluation of Mobile Apps

Measures/Rating Tools

We used 2 scales to rate the identified mobile apps. The MARS scale, assessing the quality of the mobile apps, contains 23 items across 5 subscales: engagement, functionality, aesthetics, information, and subjective quality [33]. Each item is rated on a 5-point scale from 1 (lowest quality) to 5 (highest quality). The overall app quality is assessed by calculating the mean scores of the first 4 subscales and the total mean score. The subjective quality score describes the raters’ personal liking of the app and should be reported separately if assessed. Subjective quality was not assessed in this study. The ABACUS scale, evaluating the apps’ potential to facilitate behavior change, contains 21 items across 4 subscales: knowledge and information, goals and planning, feedback and monitoring, and actions [34]. The total score is obtained by counting the number of items answered affirmatively. The MARS and the ABACUS showed good internal consistency and interrater reliability (MARS, α=.92; intraclass correlation coefficient [ICC]=0.85 and ABACUS, α=.93; ICC=0.92) [33,34].

A full evaluation of all the included apps was carried out by a first rater. A second rater independently evaluated a random sample of 30% (19/63) of the apps. Both raters were prepared for their task by completing a MARS video training tutorial [38]. To rate an app’s evidence base, as measured in MARS item 19, raters searched for randomized studies in the electronic libraries and bibliographic databases eLibrary, CyberLeninka, and Google Scholar by using the app’s name as a keyword as suggested by the MARS authors [38].

If the mobile app requested the input of demographic characteristics or consumption data, the following data were used: female gender, 30 years of age, body weight of 60 kg, height of 170 cm, and alcohol consumption on the last occasion as 200 ml of 40% vodka. If required, the maximum legal blood alcohol content was set to 0.3 ppm.

Classification of Apps and Criteria for Potential Use in SBI Programs

The identified apps were classified according to their main features. For this purpose, the following data were extracted: the app’s ability to estimate blood alcohol concentration and sobering time, its ability to record personal alcohol consumption, the presence of SBI elements, the presence of a “sobriety counter” to count the time since the last drinking occasion, and the app’s ability to support the reduction of alcohol use in a structured way. To investigate the potential of the available apps for supporting the provision of SBI in Russia, app descriptions and main features were reviewed against 2 criteria: (1) availability of AUDIT (2) whether the app provided any type of brief intervention.

XSL•FO
RenderX

Data Analysis

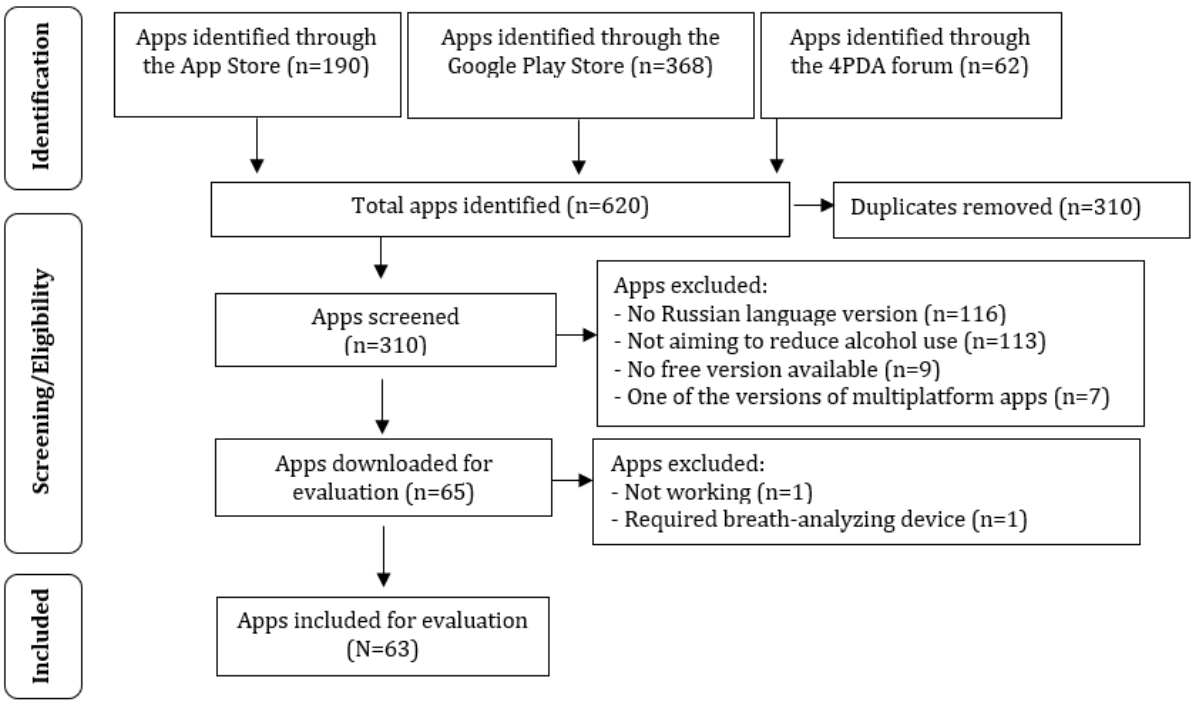
Statistical analysis and data visualization were carried out in Excel (Microsoft Excel for Office 365) and SPSS Statistics 20 (IBM Corp). Measures of interrater reliability were obtained by calculating the ICCs for all MARS and ABACUS subscales [39], using a 2-way mixed effects and average measures model with absolute agreement [40]. Descriptive analysis included total sample size, percentage, median, mean, and standard deviation.

Results

Systematic App Store Searches

A total of 620 alcohol-related apps were identified through keyword searches in the iOS App Store, Google Play Store, and 4PDA (Figure 1). After removing duplicates, 310 apps were screened against inclusion criteria, leaving 65 apps for reducing alcohol use for further download and evaluation. Among the downloaded apps, 2 had to be excluded as they did not work properly or required connection to a breathalyzer. Finally, 63 apps were included for evaluation, 51 of which were available only in the Google Play Store and 5 only in the iOS App Store. Only 7 apps were available in both stores (Multimedia Appendix 2).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart of app selection.



Overview of the Included Apps for Reducing Alcohol Use

The included apps were grouped into 6 categories according to their main feature: apps estimating blood alcohol concentration and sobering time (n=29), apps recording personal alcohol consumption (n=15), apps for SBIs (n=2), apps counting the time since the last drinking occasion (sobriety counters, n=8), apps with structured support to reduce alcohol use (n=4), and other apps for reducing alcohol use (n=5). Most apps were designed for Android systems (n=51); a minority were available for iOS (n=5) or both operation systems (n=7). A total of 19 apps were created by commercial organizations; 1 app was developed by a science center [41]. In 43 cases, no further information on the developer's legal status could be obtained. Most of the apps (53/63, 84%) had last been updated between 2016 and 2021. A minority (n=16) offered a paid upgrade version featuring the removal of advertisements and the use of additional features. According to Google Play statistics, the median number of installations was 10,000. No comparable information was provided in the iOS App Store and 4PDA. The median star rating of all apps in Google Play and iOS App Store was 3.9, based on a median of 81 ratings. Apps with structured support to reduce alcohol consumption were downloaded most often, with a median of 300,000 installations. The median star rating of the apps in this category was 4.8, based on a median of 4418 ratings.

Apps Estimating Blood Alcohol Concentration and Sobering Time

The main feature of this group of apps (n=29) was the estimation of the maximum blood alcohol concentration and sobering time. Most apps (n=14) were based on Widmark's equation [42], 2 apps used Watson's equation [43], and 13 apps provided no

XSL•FO
RenderX

information about the method of calculation. Seven apps only allowed to calculate the maximum blood alcohol concentration (n=4) or sobering time (n=3). The other 22 apps provided a combination of both mentioned features (n=16) or offered additional features such as the estimation of "no-driving" time after drinking (n=10), the estimation of the maximum alcohol consumption to sober up by a certain time (n=2), a drinking diary (n=2), or the option to unlock achievements for reducing alcohol consumption (n=1).

Apps Recording Personal Alcohol Consumption

Apps in this group (n=15) provided detailed drinking diaries (n=4), consumption calendars allowing users to indicate on what days they drunk alcohol (n=9), or both functions combined (n=2), featuring statistics of consumption per day, week, month, year. In some apps (n=4), users could calculate costs related to their alcohol use and see how much money they saved by cutting down their consumption.

Apps for SBIs

We found only 2 Russian-language apps fulfilling our criteria to potentially facilitate SBI. Both apps allowed users to complete the AUDIT. The first app provided detailed instructions for brief interventions aimed at health care professionals and a standard drink calculator allowing users to choose consumed alcoholic drinks and calculate the number of standard drinks consumed. The second app provided information on the individual level of risk and alcohol-related harm according to AUDIT results.

Apps Counting the Time Since the Last Drinking Occasion (Sobriety Counters)

We identified 8 apps in this group. Two of them consisted of a simple timer, counting the time since the last drinking occasion. The other 6 apps included additional motivation components such as a progress bar and achievements to be obtained (n=4) or inspiring citations (n=2). Two apps featured a chat where users could share their experiences. One app allowed users to consult with a medical specialist and to observe positive changes connected to alcohol abstinence in the physical appearance of the visualized avatars.

Apps With Structured Support to Reduce Alcohol Use

A total of 4 apps featured structured support to help users quit drinking or to reduce their alcohol consumption. Most of these apps provided a plan with daily tasks (n=3), a sobriety counter (n=4), a drinking diary (n=2), and motivation components. Motivation components included a progress bar and achievements to be obtained (n=3), inspiring articles or citations (n=4), daily notifications (n=4), encouraging pictures or videos (n=2), and a visualization of the positive health consequences of alcohol abstinence (n=2). One app had a community chat where users shared their experience of reducing consumption or quitting alcohol. One app provided a blood alcohol concentration calculator. Three apps allowed users to complete the AUDIT (n=2) or the Michigan Alcohol Screening Test (n=1) [44].

Other Apps for Reducing Alcohol Use

Five apps could not be assigned to any of the aforementioned categories. These included an app for audio hypnosis, an app allowing to record withdrawal symptoms, an app featuring notifications about alcohol-related harm, an app allowing to estimate the dose of consumption needed to relax, to get drunk or to have fun, and an app for counting unplanned alcohol drinking occasions after quitting drinking.

Behavior Change Techniques Featured in the Apps

The number of behavior change features provided by each app as reflected in ABACUS scores (Multimedia Appendix 2) ranged from 1 to 15 out of 21, with a mean of 5 points (SD 3.24). A great majority (54/63, 86%) of the apps requested individual baseline information and 71% (45/63) of the apps provided (individualized) user feedback. Many apps allowed the user to self-monitor their behavior (36/63, 57%) and customize or personalize certain app features (32/63, 51%). Table 1 shows the frequencies of the 21 behavioral change features evaluated in the apps. All ABACUS scores showed high interrater reliability (2-way mixed ICC=0.96; 95% CI 0.90-0.98).

The largest number of behavior change techniques was found in the categories of apps with structured support to reduce alcohol use and apps counting the time since the last drinking occasion (sobriety counters). Out of the 12 apps in these 2 groups, 9 apps featured more than 7 behavior change techniques. Apps estimating blood alcohol concentration and sobering time provided the lowest number of behavior change techniques, with an average ABACUS score of 2.38 (SD 1.37).

XSL•FO
RenderX

Table 1. Behavioral change features in the apps for reducing alcohol use (N=63).

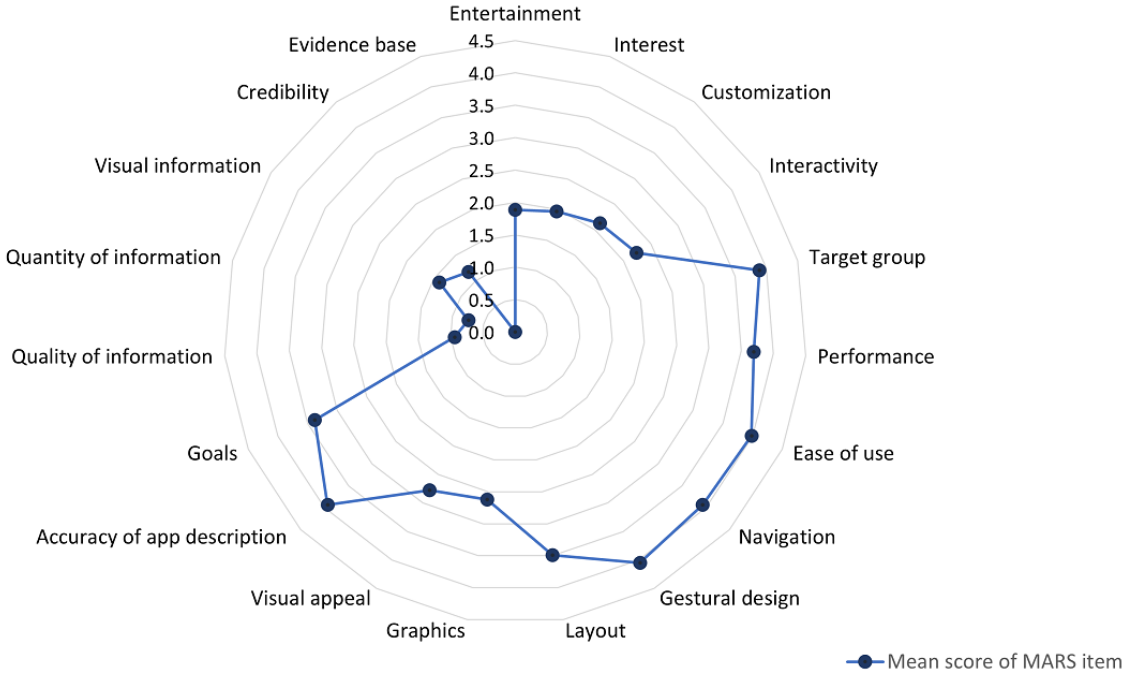
Behavioral change feature	Apps providing the feature, n (%)
Knowledge and information	
Ability to customize and personalize features	32 (51)
Consistency with national guidelines or created with expertise	3 (5)
Request for baseline information	54 (86)
Instruction on how to perform the behavior	6 (10)
Information about the consequences of continuing or discontinuing behavior	22 (35)
Goals and planning	
Request for willingness for behavior change	6 (10)
Setting of goals	7 (11)
Ability to review goals, update, and change when necessary	6 (10)
Feedback and monitoring	
Ability to quickly and easily understand the difference between current action and future goals	9 (14)
Ability to allow the user to easily self-monitor behavior	36 (57)
Ability to share behaviors with others or allow for social comparison	12 (19)
Ability to give the user feedback—either from a person or automatically	45 (71)
Ability to export data from app	4 (6)
Material or social reward or incentive	7 (11)
General encouragement	12 (19)
Actions	
Reminders or prompts or cues for activity	13 (21)
App encourages positive habit formation	7 (11)
App allows or encourages for practice or rehearsal in addition to daily activities	1 (2)
Opportunity to plan for barriers	4 (6)
Assistance with or suggest restructuring the physical or social environment	1 (2)
Assistance with distraction or avoidance	3 (5)

App Quality

Multimedia Appendix 2 shows the subscale and overall MARS scores of all evaluated apps for reducing alcohol use. Interrater reliability was high (2-way mixed ICC=0.96; 95% CI 0.91-0.98). The average overall MARS score of all reviewed apps was 2.74 (SD 0.52). MARS item 19 “evidence base” reached the lowest mean score (0.00)—no information about scientific trialing or testing of the identified app could be obtained. Other low mean scores included quantity of information (0.75, SD 1.24), quality of information (0.94, SD 1.47), credibility (1.17, SD 0.58), visual information (1.40, SD 1.75), interest (1.97, SD 0.97), and entertainment (1.89, SD 0.99). The highest scores were obtained for gestural design (4.05, SD 0.63), ease of use (3.98, SD 0.66), accuracy of app description (3.94, SD 0.56), and target group (3.89, SD 0.84). The average score of all MARS items is shown in Figure 2. A total of 30% (19/63) of the evaluated apps reached an overall MARS score of ≥3.0. The categories of apps with strong support to reduce alcohol use and apps for SBIs reached the highest overall scores (3.74 [SD 0.18] and 3.42 [SD 0.32], respectively).

XSL•FO
RenderX

Figure 2. Average score of each Mobile App Rating Scale item in Russian-language apps for reducing alcohol use (N=63). MARS: Mobile App Rating Scale.



App Potential for SBI in the Russian PHC

Out of the 63 identified mobile apps for reducing alcohol use, 4 apps contained the AUDIT, which is widely used in Russian PHC SBI. However, only 2 apps contained additional SBI elements and thus fulfilled both selection criteria. The 2 apps providing only the AUDIT ("I do not drink!" and "Sober One") contained obvious errors. The *I do not drink!* app has a "urban" translation into Russian and contains only 9 out of the 10 AUDIT questions. The *Sober One* app is potentially more attractive for SBI as it provides a brief risk assessment. However, 1 standard dose is determined as 13.7 g of pure alcohol, which does not correspond to the value officially used in Russia [17,45]. Out of the 2 apps fulfilling both selection criteria (AUDIT and Alcoholism test), the *AUDIT* app applies the SBI algorithms from the official Russian guidelines [17]. The second app, *Alcoholism test*, evaluates the individual health risk and provides additional information about the harm associated with alcohol consumption, which may motivate a conversation with the patient. However, we did not find any information on the scientific testing or trialing of the identified apps.

Discussion

Principal Findings

This study is the first systematic search and evaluation of Russian-language mobile apps for reducing alcohol use in Russia. We identified and assessed 63 eligible apps, 2 of which could potentially be used in SBIs in Russian PHC facilities after improvements in content and scientific testing. The MARS app quality scores of the evaluated apps showed good functionality, aesthetics, and ease of use. However, there is ample room for improvement, especially in the area of scientific support and evidence base; no information on scientific trialing or testing of any of the apps could be obtained. Further, ABACUS scores indicated that most apps provide only few features to facilitate human behavior change, casting doubt on their effectiveness to change alcohol consumption habits. These weaknesses seem to be common not only in Russian-language apps but in comparable apps worldwide. In a recent Australian study using both MARS and ABACUS, English-language apps for reducing alcohol use obtained similar ratings as the apps evaluated in this study [46].

The analysis of the apps' main features revealed some specific weaknesses and strong points of different app categories. Apps for calculating blood alcohol concentration and sobering time mostly used the Widmark formula developed in 1932 [42]. There are studies suggesting that this formula considerably underestimates blood alcohol concentration [47,48]. Most apps recording personal alcohol consumption and apps with structured support to reduce alcohol use featured infographics, allowing the user to quickly and visually evaluate their alcohol consumption. Some of these apps featured chats and goal-setting functions, which had an additional supportive effect. Virtual avatars with changing appearance added an element of gamification to some apps and allowed for a competitive effect between users. It is also worth highlighting the group of mobile apps with structured support to reduce alcohol use, which contained a higher number of features that facilitate changes in alcohol-related health behavior. In the future, it might be worthwhile to conduct an additional detailed analysis of these apps' content and long-term effectiveness. Users of apps for reducing alcohol use are often not aware that most apps contain unreliable, non–peer-reviewed content, which might be noneffective or even put users at risk [29,49,50]. Currently, there is no information about apps' evidence base or scientific

XSL•FO
RenderX

background available in the App Store and the Google Play Store [50].

Some international websites such as iMedicalApps offer expert comments and reviews of medical apps to patients and health care professionals [51]. Unfortunately, none of these sites are available in Russian language. SBI for alcohol consumption is not yet broadly established and implemented at the level of PHC in Russia, although decisive action was taken to change this in the past 8 years [52]. Electronic SBIs might offer greater flexibility, and depending on their mode of implementation, potentially more anonymity to avoid stigma for PHC patients [26]. Their wide availability in the App Store and, more importantly, Google Play Store as the most popular marketplace for mobile apps in Russia, may offer new opportunities to expand personalized medical care for people with alcohol-related problems [53]. The use of mobile apps to facilitate the assessment of alcohol intake as well as the level of according risk is a promising approach and requires further study, especially in a country like the Russian Federation that is committed to implement SBIs as a routine procedure.

Limitations

Searches were carried out in the iOS App Store, Google Play Store, and the 4PDA forum. These stores regularly update their content, meaning that mobile apps may become unavailable over time. Furthermore, search options such as language and region settings affect the selection and order of results, thereby reducing the reproducibility of the searches. App contents were not analyzed in detail nor did we assess the apps' potential to change human behavior in the long term. This assessment may represent an area for future research.

Conclusions

This study provides a structured overview of the main features, quality, and potential to change the alcohol-related health behavior of Russian-language apps for reducing alcohol use currently available in Russia. This overview can be used as a reference by alcohol consumers and health care professionals alike when choosing an app to facilitate the reduction of alcohol use. Although Russian-language apps for reducing alcohol use were found to be aesthetically pleasing, functional, and easy to use, most apps contained a low number of features that facilitate changes in lifestyle behavior and lacked information about scientific trialing or testing. Only 2 identified apps contained the AUDIT and additional brief intervention elements and could thus potentially be used for SBI in the Russian PHC after rigorous scientific evaluation of their effectiveness. Overall, our findings underline the need to develop evidence-based apps to mitigate alcohol consumption in Russia and elsewhere.

Authors' Contributions

AB, MN, and BG conducted the app search and evaluation. AB and VW wrote the manuscript, and BG, CFB, and MN revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PROSPERO (Prospective Register of Systematic Reviews) protocol.
[DOCX File , 26 KB - mhealth_v10i1e31058_app1.docx]

Multimedia Appendix 2
Russian-language mobile apps for reducing alcohol use.
[DOCX File , 1298 KB - mhealth_v10i1e31058_app2.docx]

References

1. Rehm J, Room R, Monteiro M, Gmel G, Graham K, Rehn N, et al. Alcohol as a risk factor for global burden of disease. Eur Addict Res 2003 Oct;9(4):157-164. [doi: 10.1159/000072222] [Medline: 12970584]
2. World Health Organization. Global Status Report on Alcohol and Health. Geneva: World Health Organization; 2018.
3. Shield K, Manthey J, Rylett M, Probst C, Wettlaufer A, Parry C, et al. National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. The Lancet Public Health 2020 Jan;5(1):e51-e61. [doi: 10.1016/s2468-2667(19)30231-2]
4. Fact sheet on the SDGs: Alcohol consumption and sustainable development. In: Sustainable Development Goals. Cham: Springer; 2020.
5. Rehm J, Gmel GE, Gmel G, Hasan OSM, Imtiaz S, Popova S, et al. The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction 2017 Jun 20;112(6):968-1001 [FREE Full text] [doi: 10.1111/add.13757] [Medline: 28220587]
6. Nemtsov A, Neufeld M, Rehm J. Are Trends in Alcohol Consumption and Cause-Specific Mortality in Russia Between 1990 and 2017 the Result of Alcohol Policy Measures? J Stud Alcohol Drugs 2019 Sep;80(5):489-498. [doi: 10.15288/jsad.2019.80.489]

XSL•FO
RenderX

7. Neufeld M, Rehm J. Alcohol consumption and mortality in Russia since 2000: are there any changes following the alcohol policy changes starting in 2006? Alcohol and Alcoholism 2013 Jan 07;48(2):222-230. [doi: 10.1093/alcalc/ags134] [Medline: 23299570]
8. Zaridze D, Brennan P, Boreham J, Boroda A, Karpov R, Lazarev A, et al. Alcohol and cause-specific mortality in Russia: a retrospective case–control study of 48 557 adult deaths. The Lancet 2009 Jun;373(9682):2201-2214. [doi: 10.1016/s0140-6736(09)61034-5]
9. Zaridze D, Lewington S, Boroda A, Scélo G, Karpov R, Lazarev A, et al. Alcohol and mortality in Russia: prospective observational study of 151 000 adults. The Lancet 2014 Apr;383(9927):1465-1473. [doi: 10.1016/s0140-6736(13)62247-3]
10. WHO launches SAFER alcohol control initiative to prevent and reduce alcohol-related death and disability. World Health Organization. 2018. URL: https://www.who.int/news/item/28-09-2018-who-launches-safer-alcohol-control-initiative-to-prevent-and-reduce-alcohol-related-death-and-disability [accessed 2021-12-13]
11. World Health Organization. Screening and brief intervention for alcohol problems in primary health care. 2006. URL: https://www.who.int/substance_abuse/publications/en/SBIFactSheet.pdf?ua=1 [accessed 2020-10-10]
12. Kaner E, Bland M, Cassidy P, Coulton S, Dale V, Deluca P, et al. Effectiveness of screening and brief alcohol intervention in primary care (SIPS trial): pragmatic cluster randomised controlled trial. BMJ 2013 Jan 09;346:e8501 [FREE Full text] [doi: 10.1136/bmj.e8501] [Medline: 23303891]
13. Kaner E, Beyer F, Muirhead C, Campbell F, Pienaar E, Bertholet N, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev 2018 Feb 24;2:CD004148-CD004131 [FREE Full text] [doi: 10.1002/14651858.CD004148.pub4] [Medline: 29476653]
14. Kaner E, Dickinson H, Beyer F, Pienaar E, Schlesinger C, Campbell F, et al. The effectiveness of brief alcohol interventions in primary care settings: a systematic review. Drug Alcohol Rev 2009 May;28(3):301-323. [doi: 10.1111/j.1465-3362.2009.00071.x] [Medline: 19489992]
15. O'Donnell A, Anderson P, Newbury-Birch D, Schulte B, Schmidt C, Reimer J, et al. The impact of brief alcohol interventions in primary healthcare: a systematic review of reviews. Alcohol and Alcoholism 2014;49(1):66-78 [FREE Full text] [doi: 10.1093/alcalc/agt170] [Medline: 24232177]
16. Garant. Methodological recommendations “Organization of medical examination and preventive medical examinations of the adult population” (approved by the Ministry of Health of the Russian Federation). Rapaht. 2013. URL: https://www.garant.ru/products/ipo/prime/doc/70229844/ [accessed 2021-12-13]
17. Garant. Methodological recommendations “Organization of preventive medical examination and clinical examination of certain groups of the adult population” (approved by the Ministry of Health on October 22). Rapaht. 2019. URL: https://www.garant.ru/products/ipo/prime/doc/72936994/ [accessed 2021-12-13]
18. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
19. World Health Organization. AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care. APA Psych Net. 2001. URL: https://doi.apa.org/doiLanding?doi=10.1037%2Ft01528-000 [accessed 2021-12-13]
20. Bunova AS, Gornyi BE, Dubovoi II, Dolgova SV, Odintsova EV, Paliy IA, et al. Primary health care workers' attitude towards to the prevention of problematic alcohol consumption. Profil. med 2017;20(5):37. [doi: 10.17116/profmed201720537-41]
21. Rosário F, Santos MI, Angus K, Pas L, Ribeiro C, Fitzgerald N. Factors influencing the implementation of screening and brief interventions for alcohol use in primary care practices: a systematic review using the COM-B system and Theoretical Domains Framework. Implement Sci 2021 Jan 07;16(1):6 [FREE Full text] [doi: 10.1186/s13012-020-01073-0] [Medline: 33413487]
22. Johnson J, Seale J. Implementing alcohol screening and brief intervention in primary care: identifying barriers, proposing solutions. Johnson and Seale Addiction Science & Clinical Practice 2014;10(S1):1-2. [doi: 10.1186/1940-0640-10-S1-A24]
23. Neufeld M, Ferreira-Borges C, Bunova A, Gornyi B, Fadeeva E, Koshkina E. Capturing Russian drinking patterns with the Alcohol Use Disorders Identification Test: an exploratory interview study in primary healthcare and narcology centers in Moscow. WHO European Office for the Prevention and Control of NCDs (unpublished manuscript) 2022:2021. [doi: 10.21203/rs.3.rs-154267/v1]
24. Kerr W, Stockwell T. Understanding standard drinks and drinking guidelines. Drug Alcohol Rev 2012 Mar;31(2):200-205 [FREE Full text] [doi: 10.1111/j.1465-3362.2011.00374.x] [Medline: 22050262]
25. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
26. Donoghue K, Patton R, Phillips T, Deluca P, Drummond C. The effectiveness of electronic screening and brief intervention for reducing levels of alcohol consumption: a systematic review and meta-analysis. J Med Internet Res 2014 Jun 02;16(6):e142 [FREE Full text] [doi: 10.2196/jmir.3193] [Medline: 24892426]

XSL•FO
RenderX

27. DIGITAL 2020 Your ultimate guide to the evolving digital world. We Are Social. URL: https://wearesocial.com/digital-2020 [accessed 2021-12-13]
28. mHealth: New Horizons for Health Through Mobile Technologies. Geneva, Switzerland: World Health Organization; 2011.
29. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
30. Bryun E, Koshkina E, Sokoltchik Y, Tetenova E, Kolgashkin A, Nadezhdin S. Mobil'nyye prilozheniya dlya lits, stradayushchikh zavisimostyami, kak element sistemy elektronnogo zdravookhraneniya (Mobile applications for addicts as part of the e-health system). Narkologiya. 2017. URL: http://www.narkotiki.ru/objects/narcology02/2017_07_spec2.pdf [accessed 2021-12-13]
31. Bryun E, Koshkina E, Tetenova E, Kolgashkin A, Nadezhdin S. Mobil'nyye prilozheniya dlya lits, ispytyvayushchikh problemy s potrebleniyem alkogolya (Mobile applications for people with alcohol consumption problems). Narkologiya. 2016. URL: http://www.narkotiki.ru/objects/narcology01/2016_04_spec_01.pdf [accessed 2021-12-13]
32. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
33. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
34. McKay FH, Slykerman S, Dunn M. The App Behavior Change Scale: Creation of a Scale to Assess the Potential of Apps to Promote Behavior Change. JMIR Mhealth Uhealth 2019 Jan 25;7(1):e11130 [FREE Full text] [doi: 10.2196/11130] [Medline: 30681967]
35. Schiavo JH. PROSPERO: An International Register of Systematic Review Protocols. Med Ref Serv Q 2019;38(2):171-180. [doi: 10.1080/02763869.2019.1588072] [Medline: 31173570]
36. Unified state automated information system. Federal Service for the Regulation of the Alcohol Market. 2021. URL: https://egais.ru/ [accessed 2021-12-13]
37. 4PDA forum. 4PDA. URL: https://4pda.ru/forum/index.php?act=idx [accessed 2021-12-13]
38. Stoyanov S, Leanne H, Kavanagh D, Tjondronegoro D, Zelenko O, Mani M. MARS Training Video. 2016. URL: https://www.youtube.com/watch?v=25vBwJQIOcE [accessed 2021-12-13]
39. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin 1979;86(2):420-428. [doi: 10.1037/0033-2909.86.2.420]
40. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023] [Medline: 22833776]
41. Bunova A, Gornyi B. Razrabotka mobil'nogo prilozheniya dlya provedeniya skrininga i profilakticheskogo konsul'tirovaniya patsiyentov s riskom pagubnogo potrebleniya alkogolya [Development of a mobile application for screening and consultation of patients with harmful alcohol consumption]. In: Profilakticheskaya Meditsina. 2019 Presented at: Non-communicable diseases and public health in Russia; September 16; Moscow p. 38.
42. Widmark EMP. Die theoretischen Grundlagen und die praktische Verwendbarkeit der gerichtlich-medizinischen Alkoholbestimmung. JAMA 1932 May 21;98(21):1834. [doi: 10.1001/jama.1932.02730470056035]
43. Watson PE, Watson ID, Batt RD. Prediction of blood alcohol concentrations in human subjects. Updating the Widmark Equation. J Stud Alcohol 1981 Jul;42(7):547-556. [doi: 10.15288/jsa.1981.42.547] [Medline: 7289599]
44. Selzer ML. The Michigan alcoholism screening test: the quest for a new diagnostic instrument. Am J Psychiatry 1971 Jun;127(12):1653-1658. [doi: 10.1176/ajp.127.12.1653] [Medline: 5565851]
45. Psikhicheskiye i povedencheskiye rasstroystva, vyzvannyye upotrebleniyem psikhoaktivnykh veshchestv. Pagubnoye (s vrednymi posledstviyami) upotrebleniye. [Mental and behavioral disorders caused by the use of psychoactive substances. Harmful (with harmful consequences) use]. Association of Narcologists of Russia. 2020. URL: https://tinyurl.com/5btkhje3 [accessed 2021-12-13]
46. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using Health and Well-Being Apps for Behavior Change: A Systematic Search and Rating of Apps. JMIR Mhealth Uhealth 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: 10.2196/11926] [Medline: 31274112]
47. Brouwer I. The Widmark formula for alcohol quantification. SADJ 2004 Nov;59(10):427-428. [Medline: 15726763]
48. Friel PN, Logan BK, Baer J. An Evaluation of the Reliability of Widmark Calculations Based on Breath Alcohol Measurements. J. Forensic Sci 1995 Jan 01;40(1):13767J. [doi: 10.1520/jfs13767j]
49. Capras R, Bolboaca S. An evaluation of free medical applications for android smartphones. Applied Medical Informatics (3-4). 2016. URL: https://www.researchgate.net/publication/338966234_An_Evaluation_of_Free_Medical_Applications_for_Android_Smartphones [accessed 2021-12-13]
50. Visser BJ, Buijink AWG. Need to peer-review medical applications for smart phones. J Telemed Telecare 2012 Mar;18(2):124. [doi: 10.1258/jtt.2011.110205] [Medline: 22289574]
51. Husain I, Misra S, Wodajo F. iMedicalApps. 2019. URL: https://www.imedicalapps.com/ [accessed 2021-12-13]

XSL•FO
RenderX

52. Neufeld M, Bunova A, Gornyi B, Ferreira-Borges C, Gerber A, Khaltourina D, et al. Russia's National Concept to Reduce Alcohol Abuse and Alcohol-Dependence in the Population 2010-2020: Which Policy Targets Have Been Achieved? Int J Environ Res Public Health 2020 Nov 09;17(21):8270 [FREE Full text] [doi: 10.3390/ijerph17218270] [Medline: 33182377]
53. Russia's mobile app economy: A new app superpower in the making? AppAnnie.com. URL: https://www.appannie.com/en/insights/market-data/russias-mobile-app-economy/ [accessed 2021-12-13]

Abbreviations

ABACUS: App Behavior Change Scale
AUDIT: Alcohol Use Disorders Identification Test
AUDIT-C: Alcohol Use Disorders Identification Test-Consumption
ICC: intraclass correlation coefficient
iOS: iPhone operating system
MARS: Mobile App Rating Scale
PHC: primary health care
PROSPERO: Prospective Register of Systematic Reviews
SBI: screening and brief intervention
WHO: World Health Organization

Edited by L Buis; submitted 08.06.21; peer-reviewed by I Shubina, H Goodfellow; comments to author 30.07.21; revised version received 09.08.21; accepted 29.10.21; published 10.01.22.

Please cite as:
Bunova A, Wiemker V, Gornyi B, Ferreira-Borges C, Neufeld M
Russian-Language Mobile Apps for Reducing Alcohol Use: Systematic Search and Evaluation
JMIR Mhealth Uhealth 2022;10(1):e31058
URL: https://mhealth.jmir.org/2022/1/e31058
doi:10.2196/31058
PMID:35006083

©Anna Bunova, Veronika Wiemker, Boris Gornyi, Carina Ferreira-Borges, Maria Neufeld. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 10.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prioritization of Quality Principles for Health Apps Using the Kano Model: Survey Study

Christin Malinka[1*]; Ute von Jan[1*], Dr rer biol hum; Urs-Vito Albrecht[1,2*], MD, DrPH, PhD

[1]Peter L Reichertz Institute for Medical Informatics, TU Braunschweig and Hannover Medical School, Hannover, Germany

[2]Department of Digital Medicine, Medical Faculty OWL, Bielefeld University, Bielefeld, Germany

[*]all authors contributed equally

Corresponding Author:
Ute von Jan, Dr rer biol hum
Peter L Reichertz Institute for Medical Informatics
TU Braunschweig and Hannover Medical School
Carl-Neuberg-Str. 1
Hannover, 30625
Germany
Phone: 49 511 532 ext 4412
Fax: 49 511 532 19335
Email: ute.von.jan@plri.de

Abstract

Background: Health apps are often used without adequately taking aspects related to their quality under consideration. This may partially be due to inadequate awareness about necessary criteria and how to prioritize them when evaluating an app.

Objective: The aim of this study was to introduce a method for prioritizing quality attributes in the mobile health context. To this end, physicians were asked about their assessment of nine app quality principles relevant in health contexts and their responses were used as a basis for designing a method for app prioritization. Ultimately, the goal was to aid in making better use of limited resources (eg, time) by assisting with the decision as to the specific quality principles that deserve priority in everyday medical practice and those that can be given lower priority, even in cases where the overall principles are rated similarly.

Methods: A total of 9503 members of two German professional societies in the field of orthopedics were invited by email to participate in an anonymous online survey over a 1-month period. Participants were asked to rate a set of nine app quality principles using a Kano survey with functional and dysfunctional (ie, positively and negatively worded) questions. The evaluation was based on the work of Kano (baseline), supplemented by a self-designed approach.

Results: Among the 9503 invited members, 382 completed relevant parts of the survey (return rate of 4.02%). These participants were equally and randomly assigned to two groups (test group and validation group, n=191 each). Demographic characteristics did not significantly differ between groups (all $P>.05$). Participants were predominantly male (328/382, 85.9%) and older than 40 years (290/382, 75.9%). Given similar ratings, common evaluation strategies for Kano surveys did not allow for conclusive prioritization of the principles, and the same was true when using the more elaborate approach of satisfaction and dissatisfaction indices following the work of Timko. Therefore, an extended, so-called “in-line-of-sight” method was developed and applied for this evaluation. Modified from the Timko method, this approach is based on a “point of view” (POV) metric, which generates a ranking coefficient. Although the principles were previously almost exclusively rated as *must-be* (with the exception of resource efficiency), which was not conducive to their prioritization, the new method applied from the *must-be* POV resulted in identical rankings for the test and validation groups: (1) legal conformity, (2) content validity, (3) risk adequacy, (4) practicality, (5) ethical soundness, (6) usability, (7) transparency, (8) technical adequacy, and (9) resource efficiency.

Conclusions: Established survey methodologies based on the work of Kano predominantly seek to categorize the attributes to be evaluated. The methodology presented here is an interesting option for prioritization, and enables focusing on the most important criteria, thus saving valuable time when reviewing apps for use in the medical field, even with otherwise largely similar categorization results. The extent to which this approach is applicable beyond the scenario presented herein requires further investigation.

(JMIR Mhealth Uhealth 2022;10(1):e26563) doi:10.2196/26563

XSL•FO
RenderX

KEYWORDS

Kano; quality principles; mobile apps; physicians; surveys and questionnaires; evaluation studies; mHealth; health apps

Introduction

Background

Independent of their proficiency with apps and the respective usage contexts, users are often unfamiliar with the intricacies of the specific aspects that are essential for recognizing an app's quality. Even apps covering health contexts are often marketed without having been evaluated by experts, and with only minimally relevant and reliant information being provided (eg, regarding scientific studies [1,2]). Thus, for end users, making an informed decision about whether or not to use an app is not an easy task, independent of whether they are health care professionals, patients with chronic conditions, or even laypeople with a more generic interest in health apps.

There are numerous, more or less elaborate, tools, norms, and lists of quality criteria that either target developers or aim at aiding those interested in an app in their decision process (eg, [3-8]), and many of the aspects they cover overlap. However, even if interested parties are aware of these approaches, if a quick assessment is desired, these approaches may sometimes be seen as going too far or being too complex. Both the paucity of readily available information or expert assessments [1,9] in identifying apps that can be recognized as trustworthy, as well as the difficulty in identifying suitable criteria for an initial and independent assessment, can mean that apps often fail to realize the potential attributed to them for medical care and prevention [10-13]. Checklists that interested users may apply to apps (eg, [4,14,15]) often target careful curation of a list of apps for later use, but may be too extensive for practical application and quick assessments in everyday medical practice. It may therefore be helpful to develop and apply a process for identifying a subset of criteria or quality principles listed in such tools considered to be particularly relevant for a specific target group, which may be achieved by means of prioritization.

As a foundation for this study, we used nine basic quality principles for health apps that were previously compiled [16,17] and evaluated [18,19] in a multistep process: (1) practicality, (2) risk adequacy, (3) ethical soundness, (4) legal conformity, (5) content validity, (6) technical adequacy, (7) usability, (8) resource efficiency, and (9) transparency. Participants in both of the aforementioned evaluation studies were first requested to provide initial assessments regarding the perceived relevance of these principles. They were then provided with applied app store descriptions and asked to determine whether they deemed the textual information sufficient to satisfy the above principles. Subsequently, they were asked to apply 25 questions operationalizing the nine principles to the same store descriptions. Between each of the steps, they were asked whether or not they would consider using the respective app based on the available information. During the course of these studies [18,19], as participants familiarized themselves with the quality criteria, they were able to make a more confident, but increasingly critical, assessments of the apps based on the available information.

These previous studies with medical students [18] and members of the German Society for Internal Medicine [19] showed that the participants predominantly perceived all nine of the above quality principles as important. For both studies, the data were evaluated using two (randomly assigned and equally sized) test and validation groups [18,19]. Although there were no significant differences in the answers obtained for the nine principles between the two groups, solely based on assigned relevance, rankings (and thus any prioritizations based on them) would have differed [18] between the groups as well as between the two studies. Apart from slightly lower relevance ratings for resource efficiency, all other quality principles were seen as either "important" or "very important"; however, owing to their closeness with respect to the ratings, any order of the principles based on these ratings seemed to have been influenced by statistical noise rather than sound calculations. Nevertheless, in both of the aforementioned studies [18,19], some participants expressed fear that the application of even these few principles would be too time-consuming for use in an everyday care context. As even the relevance-related questions for the criteria that were asked in these studies did not allow for their ranking, we therefore aimed to establish a method that would meet the demand for a better focus on quality aspects for mobile health (mHealth) apps that would be perceived as particularly relevant in the community.

Hypotheses

We hypothesized that methods established to assess product attributes in marketing-related research might also be suitable for categorizing quality attributes for mHealth apps. We tested this hypothesis based on an exemplary Kano survey related to the nine aforementioned quality principles. In this type of survey, questions are implemented based on a model developed by Noriaki Kano in the 1970s and 1980s. The "Kano model" is often used in the context of marketing or for refining products, specifically with regard to customer satisfaction with a product's features in mind [20]. As Kano noted, there need not be a linear relationship between satisfaction or dissatisfaction and the fulfillment of a need [21]; thus, to be able to nevertheless assess a product, he proposed using so-called "functional" and "dysfunctional" questions that not only assess a participant's opinion about a feature being available but also about it not being provided.

On its own, if successful at all, such a Kano survey–based categorization can only provide a rough prioritization at best, based on ranking the categories according to their fitness for the question at hand. As this approach may fail in cases where the attributes under consideration are rated similarly, we established our second hypothesis that it should be possible to nevertheless prioritize the product attributes studied (in our case, the nine quality principles) by developing and applying an extended method on the basis of the data collected.

Objectives

This study builds upon the foundation laid by previous studies in the health app quality context. This work was motivated by

XSL•FO
RenderX

interest to find and apply a method that helps to more finely differentiate between a chosen set of quality attributes to be used in such a setting. As indicated above, although there are a variety of tools for this task or lists of quality principles for different app types in the mHealth domain, there are voices lamenting that despite these tools being academically sound, applying them in a real-world setting or for a large number of apps may be too tedious [22].

In our evaluation, the proposed method was applied to the nine predefined health app quality principles to determine whether it is feasible to determine an adequate and stable ranking of such criteria to be used for prioritization in facilitating app assessments should the need arise.

Basic Design of the Study

Our approach is based on a group of popular techniques for classifying quality attributes that are often used in decision-making processes in the areas of marketing, management, or even a product's design phase [23] if a decision is to be made about which (planned or existing) attributes of a product elicit customer satisfaction (and should thus be used or further investigated for a product) or dissatisfaction (making them superfluous or even counterproductive for the product's success). Following this line of thought, we used a survey design based on Kano's model of attractive quality for classifying quality attributes (originally published in Japanese [20] and subsequently in English [24]), and applied various more elaborate evaluation techniques as specified in the literature (eg, those proposed by Timko as cited in Berger et al [25]) to the acquired data.

Using the Kano survey data and available evaluation methods, it may be conceivable to find sufficiently differing categorizations of the quality principles that allow for selecting a particularly relevant subset of principles based on their assigned (Kano or derived) category, whereas principles in lower-ranking or less-desirable categories are treated as deferred or are even removed from further consideration. As applied to the nine quality principles, we suspected that even if the principles are largely seen as similarly important, some might be viewed as more attractive, essential, or indifferent than others. Based on a per-category ranking (depending on the perceived relevance of the categories for the use case), we deemed it possible to determine at least a partial prioritization.

As the first idea was unfortunately quickly disproved due to the largely similar categorizations of the nine principles based on the acquired survey data, as a second approach, we tried to better take into account to what degree a product's attributes, or in our case the app quality principles, contribute to (customer) satisfaction or dissatisfaction, specifically based on the work proposed by Timko in Berger et al [25]. Our assumption was that by appropriately taking both the numeric values for satisfaction as well as dissatisfaction into account, it should be possible to determine a numeric representation in the form of a ranking coefficient (eg, using a ratio of the two values or similar approaches) that could lay the foundation for finding a relatively stable means for prioritization of app quality principles based on this value.

Methods

Data Acquisition

Implementation

Data collection for the study took place in the form of an anonymous and data protection–compliant online survey, implemented using the SoSci Survey [26] installation provided at Hannover Medical School. The survey was open for 1 month (between December 2, 2019, and January 2, 2020), and using the mailing lists of both the German Society for Orthopedics and Trauma Surgery (DGOU) and the Orthopedics and Trauma Surgery Professional Association (BVOU); a total of 9503 members of these societies were invited to participate.

Prior to sending the survey invitation, the study was reviewed by the Ethics Committee of Hannover Medical School (application number 8746_BO_K_2019). In the vote dated November 4, 2019, no ethical or legal objections were raised.

Structure of the Survey

The actual survey itself was conducted in two parts. The first part contained questions about the German Digital Healthcare Act (DVG [27]) that, at the time of the survey, had recently been ratified. Participants were presented with questions about their familiarity with this act, their opinions about its coverage, and whether they were at all considering making use of the possibility to prescribe health apps based on the processes specified in the DVG. The data corresponding to this part of the survey were previously evaluated and published [28].

To acquire demographic data, those responding to the survey were asked questions related to age and gender, as well as about their work history and environment (how long they had been working; their current function; and whether they were working in private practice, at a clinic, or another institution). To allow a basic assessment about their familiarity with mHealth, they were also asked about their private and work-related usage of mHealth apps, and whether any patients asked them either about specific health apps or about a recommendation for a health app. However, the demographic data are only presented to describe the participating physicians. Apart from exemplary calculations given in the Discussion, these data were not part of the analyses presented in this paper.

The work presented herein specifically deals with the second part of the survey. As mentioned in the Introduction, a predefined set of nine quality principles (practicality, risk adequacy, ethical soundness, legal conformity, content validity, technical adequacy, usability, resource efficiency, and transparency) was employed as a basis for the evaluation. The set of quality principles has previously been published [16, 17] along with their evaluations [18,19].

In the context of the work presented here, following Kano's method, for each of the nine quality principles, the participants were presented with a set of so-called functional and dysfunctional questions (see Table 1). Answer options for both types of questions were "I would be very pleased," "I'd expect this," "I don't care," "I could accept that," and "That would really bother me."

XSL•FO
RenderX

Table 1. Quality principles with the corresponding questions (translated from the original German-language version) for functional and dysfunctional aspects, as required by the Kano model.

Principle	Functional question	Dysfunctional question
Practicality	What would you say if apps could be used for the intended purpose?	What would you say if apps could not be used for the intended purpose?
Risk adequacy	What would you say if apps did not pose a disproportionate health, social, or economic risk to users?	What would you say if apps posed disproportionate health, social, or economic risks to users?
Ethical soundness	What would you say if discrimination and stigmatization were avoided when developing, offering, and using apps?	What would you say if discrimination or stigmatization were not avoided when developing, offering, operating, and using apps?
Legal conformity	What would you say if apps were compliant with data protection regulations as well as professional and health regulations?	What would you say if apps failed to comply with data protection, professional, or health regulations?
Content validity	What would you say if the content used in apps was valid and trustworthy?	What would you say if the content used in apps was not valid or not trustworthy?
Technical adequacy	What would you say if apps were easy to maintain and could be used independent of a specific platform?	What would you say if apps were hard to maintain or could not be used independent of a specific platform?
Usability	What would you say if apps were designed and implemented according to the requirements of the target group(s)?	What would you say if apps were not designed and implemented to meet the needs of the target group(s)?
Resource efficiency	What would you say if apps were to use resources such as battery and computing power efficiently?	What would you say if apps made only inefficient use of resources such as battery or computing power?
Transparency	What would you say if apps provided transparent information about inherent quality features?	What would you say if apps did not provide transparent information about inherent quality characteristics?

In addition to the functional and dysfunctional questions, the participants were also asked to rate the perceived relevance for each of the nine principles (Table 2). In this case, answers could be given using a 5-point scale: "very important," "important," "neutral," "less important," and "unimportant."

For each quality principle, the "functional" question was always presented first, followed by the "dysfunctional" question, and that for relevance. However, for each participant, the order in which the questions were shown was randomly assigned to alleviate bias based on an attribute's position in the list.

Table 2. Questions regarding the relevance for each of the nine quality principles (translated from the original German version).

Principle	Perceived relevance
Practicality	How important is it to you that apps can be used for the intended purpose?
Risk adequacy	How important is it to you that apps are low risk in terms of health, social, or economic risks?
Ethical soundness	How important is it to you to avoid discrimination and stigmatization when developing, offering, operating, and using apps?
Legal conformity	How important is it to you that data protection, professional, and health regulations are respected in apps?
Content validity	How important is the validity and trustworthiness of the health-related content presented and used in an app to you?
Technical adequacy	How important are easy maintainability and platform-independent or cross-platform usability of apps to you?
Usability	How important is the target group–oriented design and operation of apps to you?
Resource efficiency	How important to you is the efficient use of resources through apps, for example in terms of battery and computing power?
Transparency	How important is it to you that apps provide transparent information about inherent quality features?

Categorization of Answers According to Kano

Using the Kano model, based on the answers given for both functional and dysfunctional questions (see Table 3), a product's features can be categorized as *attractive (A)*, if its presence leads to satisfaction but there is no (additional) dissatisfaction if it is missing [25]; *must-be (M)*, if the respective feature is deemed essential (ie, if it does not improve satisfaction if available, but leads to extreme dissatisfaction if missing) [29]; *one-dimensional (O)*, also referred to as the *performance (P)* category in the literature, if both availability and lack of the feature cause satisfaction and dissatisfaction, respectively [25], thus representing a feature that customers explicitly demand; *indifferent (I)*, if the feature (or the lack thereof) influences neither satisfaction nor dissatisfaction, thus being ideal for elimination if a reduction in overhead is desired [30]; *reverse (R)*, if dissatisfaction is caused if the feature is available and satisfaction if it is missing; and *questionable (Q)* if the answers given to the functional and dysfunctional questions are in contradiction [25] (eg, if both answers are specified as "I would be very pleased").

XSL•FO
RenderX

Table 3. Assignment of answers to various categories to both functional and dysfunctional questions (based on [25]) and representation of answer pairs where one or both answers are missing.

Answers to functional questions	Answers to dysfunctional questions					
	I would be very pleased	I'd expect this	I don't care	I could accept that	That would really bother me	No answer given
I would be very pleased	Q[a]	A[b]	A	A	P[c]	—[d]
I'd expect this	R[e]	Q	I[f]	I	M[g]	—
I don't care	R	I	I	I	M	—
I could accept that	R	I	I	Q	M	—
That would really bother me	R	R	R	R	Q	—
No answer given	—	—	—	—	—	—

[a]Q: questionable.
[b]A: attractive.
[c]P: performance (one-dimensional).
[d]Not applicable.
[e]R: reverse.
[f]I: indifferent.
[g]M: must-be.

Both *the reverse and questionable* categories may, for example, be due to inadequate wording of the questions employed in the survey or side effects from other (not necessarily easily explainable) factors that impact the answers. Especially for the questionable category, the answers given may also indicate that a participant was (for whatever reason) unwilling to answer in a sensible manner.

Evaluation Strategies

For each of the nine quality principles, the answers provided by the participants for the functional and dysfunctional question pairs were then categorized based on Table 3, and the frequency that each category was assigned to each attribute was calculated. These counts were then used for further evaluation. As described previously [25,31], there are several strategies that can be applied for this task.

One approach is to determine the category for a feature based on its greatest frequency. Alternatively, an if-then–based approach can be adopted: if $(P+A+M)>(I+R+Q)$, the category that corresponds to the maximum count for *performance*, *attractive*, or *must-be* is used; however, if $(P+A+M)<(I+R+Q)$, the category corresponding to the maximum of *indifferent*, *reverse*, or *questionable* as the category assigned to the feature under consideration is used.

Both of these approaches work best if those surveyed are somewhat consistent in their answers for a specific feature, or at least show a clear tendency toward a specific category for that feature. However, these approaches do not work quite as well if the responses are distributed more evenly across several categories such as *attractive*, *performance*, *must-be*, and *indifferent*. Moreover, if different features elicit similar responses, it may be difficult to discriminate between them. This may hamper the usefulness of the approach in the context of categorization.

Timko (cited in [25]) proposed an additional method, as he noted that based on the aforementioned mode statistic, the results may seem somewhat skewed. For example, for two features with only *attractive* and *indifferent* ratings, albeit one with a 90-to-10 *attractive*-to-*indifferen*t ratio and the other with only a 60-to-40 *attractive*-to-*indifferent* ratio, the assigned category will be *attractive* for both. Thus, a third method tries to alleviate these disadvantages.

This method uses the previously obtained counts to calculate two distinct values: one representing the relative value of meeting a customer requirement (namely, "what if we're better" in contrast to a competitor) and the other representing the relative cost of not meeting the customer requirement (ie, worse than the competition). The two values, as defined in Berger et al [25], are calculated as follows:

$$\text{Better} = (A+P)/(A+O+M+I), \text{ with } 0 \leq \text{Better} \leq 1$$

$$\text{Worse} = -(O+M)/(A+O+M+I), \text{ with } -1 \leq \text{Worse} \leq 0$$

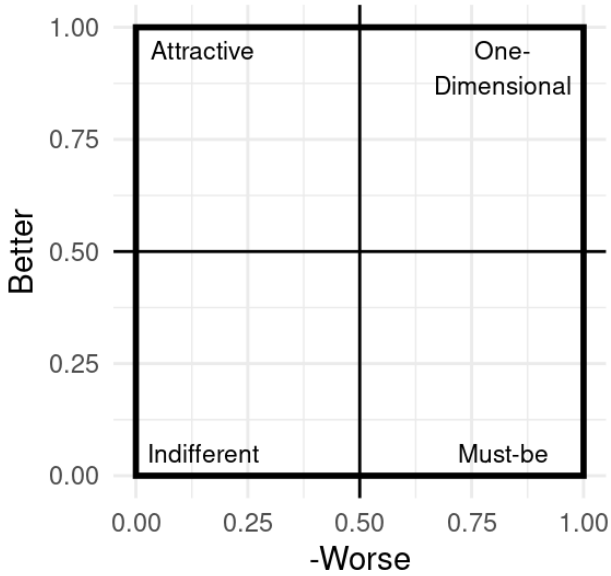
On average, satisfaction will increase for *attractive* and *one-dimensional* (*performance*) attributes, which is why, in the literature, "Better" is also often denoted as the satisfaction index [32,33], and satisfaction decreases if *one-dimensional* and *must-be* elements are not adequately represented. For this reason, *"Worse"* is often called the dissatisfaction index [32,33]. Both *questionable* as well as *reverse* answers are ignored in Timko's approach, but nevertheless, the calculations do respect a possible spread of the attributes under consideration over the different categories.

The Worse-Better pairing for calculated attributes can be plotted on a two-dimensional and easy-to-interpret graph. Commonly, the values for each attribute are additionally multiplied by the average relevance the participants assign to each attribute to improve discrimination between value pairs for features located in direct vicinity to each other. According to Timko, when deciding which attributes to keep or to omit, one should choose

XSL•FO
RenderX

those for which satisfaction (ie, the Better score) is higher, since they add more to customer satisfaction, whereas on the Worse axis, one should aim for more negative values, as they prevent dissatisfaction [25] (Figure 1).

Figure 1. Two-dimensional representation of Worse-Better pairings for the Kano quality categories [25]. For easier interpretation, Worse is shown with its absolute value.



Designing an Improved Methodology for Prioritization

Discussions among the authors led to the conclusion that established methods such as those described above were suffering from only being able to assign broadly defined categories to the attributes under consideration, without allowing for a more granular consideration that actually respects the relative location of the attributes under consideration. This is particularly relevant when the attributes to be compared (represented by their Worse and Better coordinates) are (predominantly) located in one of the four quadrants and are therefore assigned to the same category (ie, *indifferent*, *must-be*, *attractive*, or *one-dimensional*). With this in mind, we designed an "in-line-of-sight" method that allows for rankings depending on different points of view on the coordinate system.

This new approach makes it possible to establish a reference to the proximity of an attribute's (or quality principle's) coordinate points to the respective outermost corner (corresponding to the point most clearly representing the quadrant), and further respects their relative positions for obtaining the ranking.

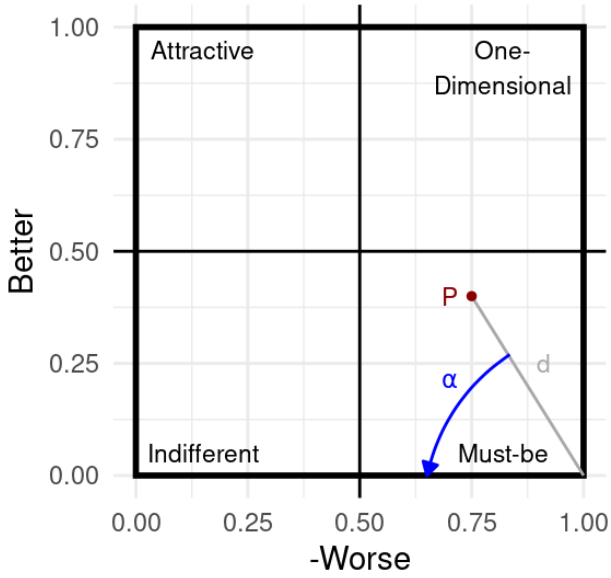
This approach will now be explained in more detail by way of an example, using the *must-be* quadrant as a point of reference. Starting from the outermost point of this category, denoted by the coordinates ($Worse_I$=–1, $Better_I$=0), for each attribute (or quality principle), the Euclidean distance between this point and the respective coordinate is first calculated. An increasing distance to the *must-be* corner represents a greater proximity to one of the three other categories (and is, as such, less desirable).

For further improved differentiation between quality principles, even in the case of (almost) identical Euclidean distances, an angle is then determined based on the chosen secondary ranking strategy. In our example (and all further calculations shown in this paper), we decided to prefer points with less pronounced Worse values (ie, those that have less potential for causing dissatisfaction according to Timko). For this purpose, we chose to calculate an offset based on the angle (denoted by α) between the x-axis of the coordinate system and the line defined by the corner point's coordinate p=(–1,0) as well as the respective quality principle's $q=(-Worse_I, Better_I)$ coordinate (see Figure 2). As α is only supposed to aid with differentiation between points with similar distance values, it needs to be rescaled to an appropriate value range. First, α is divided by the maximum possible angle (ie, 90°) and then multiplied with 0.05×2≈0 (representing 5% of the maximum possible distance of the square root of 2 in the coordinate system). The distance and adapted angle value are then summarized (hereinafter referred to as the ranking coefficient f), and the resulting value for f is then used for ranking the quality principles according to ascending order as follows:

For simplification, as the plots use an inverted x-axis for representing the Worse value, all statements (as well as the angle calculations) concerning the left- or right-hand location of any point or axis mentioned in relation to the coordinate system refer to this inverted plot. For the other three quadrants, if necessary, rankings may be performed in a similar manner.

XSL•FO
RenderX

Figure 2. Angle (α) and distance (d) for a point (P) located in the must-be corner, as employed in the in-line-of-sight method (seen from the must-be corner).



Statistics Tools

The R language and environment for statistical computing, version 4.0, was used for all evaluations, along with accompanying packages such as dplyr, ggplot2, arsenal, and others [34-36].

Results

Data

Of those who answered our survey, only 382 actually completed all of its parts, and were thus included in the evaluation presented here. This corresponds to a return rate of 4.02% of the 9503 potential participants.

Using the sample_frac function provided by the dplyr package [34], the available participants were randomly assigned to the test (group A, n=191) and validation (group B, n=191) groups.

Baseline Demographics of the Participants

To rule out differences between the two groups due to demographic factors, these were first compared. There were no statistically significant differences between the groups with respect to baseline demographics (*P*>.05 for all factors, see Table 4). Overall, the participants were predominantly male and older than 40 years (290/382, 75.9%). In line with the age structure, over three-quarters of the participants had a work experience of more than 10 years (288/328, 75.4%; excluding retirees, 19/328, 5.8%) and were working in higher-level functions (attendings, chiefs, or specialists in private practice; 284/382, 74.3%). The majority of participants worked in a hospital setting (acute care or university hospital; 232/382, 60.7%). As we had only surveyed members of two German orthopedic societies, the proportion of those who were not active in Germany was low, as expected (10/382, 2.6%).

Although the participants overwhelmingly stated that they were highly interested or interested in digital technology (316/382, 82.7%), this was not mirrored by the proportion of those admitting to app use in private or work settings. Only slightly over one-fifth of those participating had already been asked by patients about a specific app or about recommending an app (see Table 4 for full data).

XSL•FO
RenderX

Table 4. Base demographics for all participants and for those assigned to the test group (A) and validation group (B).

Characteristic	Group A (n=191), n (%)	Group B (n=191), n (%)	Total (N=382), n (%)	*P* value[a]
Age group (years)				.87
21-30	9 (4.7)	7 (3.7)	16 (4.2)	
31-40	34 (17.8)	42 (22.0)	76 (19.9)	
41-50	46 (24.1)	44 (23.0)	90 (23.6)	
51-60	62 (32.5)	59 (30.9)	121 (31.7)	
>60	40 (20.9)	39 (20.4)	79 (20.7)	
Gender				.38
Female	24 (12.6)	30 (15.7)	54 (14.1)	
Male	167 (87.4)	161 (84.3)	328 (85.9)	
Work experience				.93
Not yet working	2 (1.0)	1 (0.5)	3 (0.8)	
<1 year	2 (1.0)	2 (1.0)	4 (1.0)	
1-5 years	10 (5.2)	14 (7.3)	24 (6.3)	
6-10 years	19 (9.9)	25 (13.1)	44 (11.5)	
11-20 years	50 (26.2)	44 (23.0)	94 (24.6)	
21-30 years	54 (28.3)	50 (26.2)	104 (27.2)	
>30 years	44 (23.0)	46 (24.1)	90 (23.6)	
Retired	10 (5.2)	9 (4.7)	19 (5.0)	
Professional level				.75
Student	1 (0.5)	0 (0.0)	1 (0.3)	
In training/resident	23 (12.0)	25 (13.1)	48 (12.6)	
Attending	60 (31.4)	52 (27.2)	112 (29.3)	
Chief	38 (19.9)	39 (20.4)	77 (20.2)	
Specialist (private practice)	47 (24.6)	48 (25.1)	95 (24.9)	
Other	21 (11.0)	27 (14.1)	48 (12.6)	
Not answered	1 (0.5)	0 (0.0)	1 (0.3)	
Work setting				.49
Acute care: standard care level	63 (33.0)	50 (26.2)	113 (29.6)	
Acute care: maximum care level	32 (16.8)	37 (19.4)	69 (18.1)	
University hospital	21 (11.0)	29 (15.2)	50 (13.1)	
Rehabilitation center	8 (4.2)	7 (3.7)	15 (3.9)	
Medical care center	6 (3.1)	9 (4.7)	15 (3.9)	
Private practice	40 (20.9)	44 (23.0)	84 (22.0)	
Other	21 (11.0)	14 (7.3)	35 (9.2)	
Not answered	0 (0.0)	1 (0.5)	1 (0.3)	
Geographic location[b]				.26
Germany	187 (98.9)	183 (95.8)	370 (97.4)	
Austria	0 (0.0)	2 (1.0)	2 (0.5)	
Switzerland	2 (1.1)	3 (1.6)	5 (1.3)	
Other: European Union	0 (0.0)	2 (1.0)	2 (0.5)	
Other: not yet listed	0 (0.0)	1 (0.5)	1 (0.3)	

XSL•FO
RenderX

Characteristic	Group A (n=191), n (%)	Group B (n=191), n (%)	Total (N=382), n (%)	*P* value[a]
Interest in digital technology				.71
Highly interested	76 (39.8)	81 (42.4)	157 (41.1)	
Interested	84 (44.0)	75 (39.3)	159 (41.6)	
Neutral	19 (9.9)	25 (13.1)	44 (11.5)	
Less interested	8 (4.2)	8 (4.2)	16 (4.2)	
Not interested	4 (2.1)	2 (1.0)	6 (1.6)	
Uses apps in private settings				.92
Yes	69 (36.1)	70 (36.6)	139 (36.4)	
No	122 (63.9)	121 (63.4)	243 (63.6)	
Uses apps for work				.29
Yes	63 (33.0)	73 (38.2)	136 (35.6)	
No	128 (67.0)	118 (61.8)	246 (64.4)	
Been asked about an app/recommendation				>.99
Yes	43 (22.5)	43 (22.5)	86 (22.5)	
No	148 (77.5)	148 (77.5)	296 (7.5)	

[a]Pearson χ^2 test.

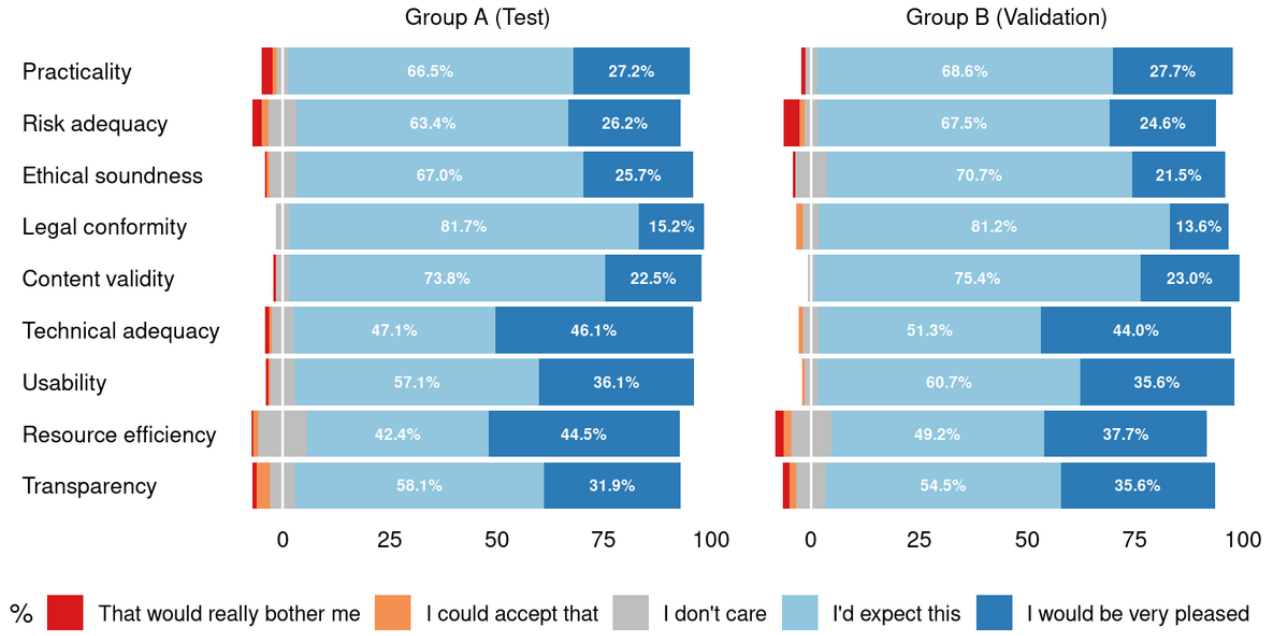
[b]Not answered: group A, n=2.

Data Evaluation

Descriptive Evaluation of the Survey Results

Similar to the participants' demographics, in the Kano-based questionnaire, there were no statistically significant differences between the training and validation groups with respect to answers given for the functional and dysfunctional questions, as well as the perceived relevance for the nine app quality criteria (see Figures 3, 4, and 5; for more detailed counts, proportions, and *P* values for the available answers, see Multimedia Appendix 1, Tables S1-S3).

Figure 3. Distribution of answers for the functional questions. For legibility reasons, smaller values are not printed (see Multimedia Appendix 1 for the complete list of values).



XSL•FO
RenderX

Figure 4. Distribution of answers for the dysfunctional questions. For legibility reasons, smaller values are not printed (see Multimedia Appendix 1 for the complete list of values).

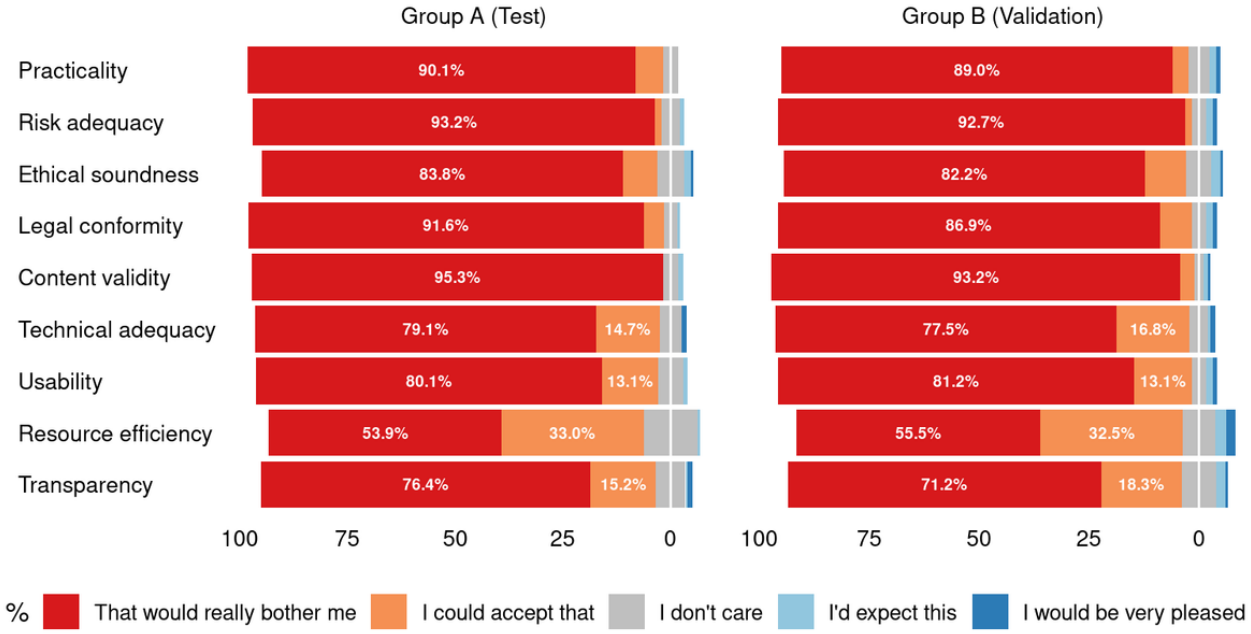
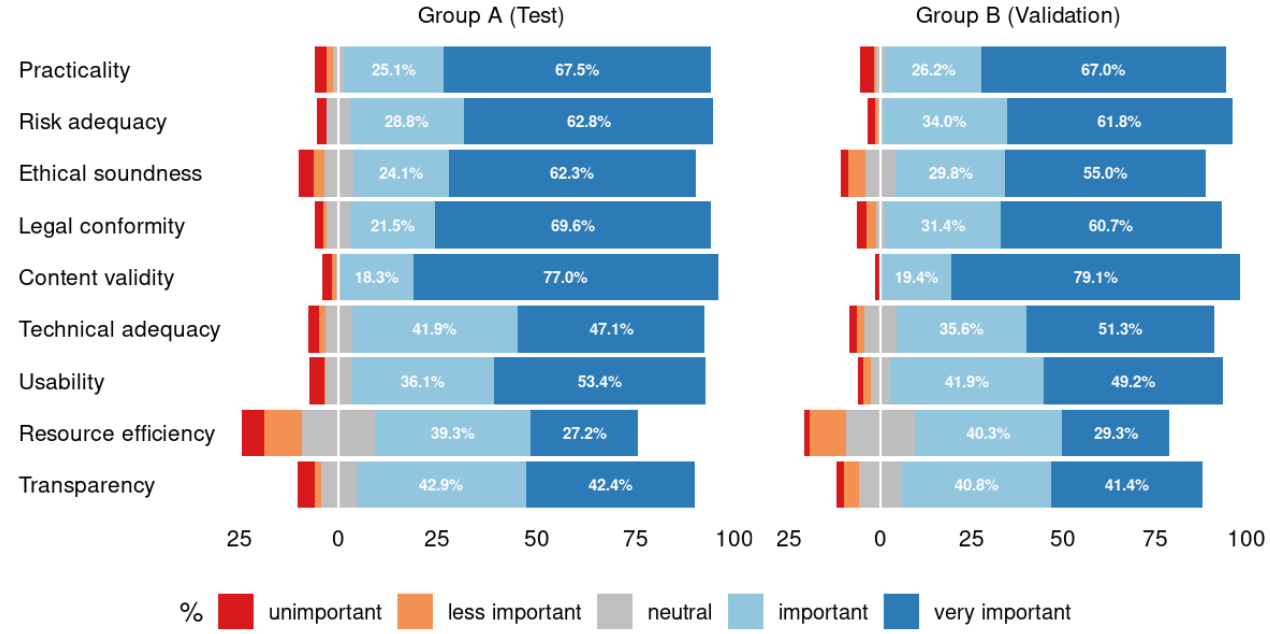


Figure 5. Ratings for relevance of the nine quality principles, as perceived by the participants. For legibility reasons, smaller values are not printed (see Multimedia Appendix 1 for the complete list of values).



Categorization According to Kano

Using Kano's basic evaluation described in the "Evaluation Strategies Applied" subsection within the Methods, namely choosing the category with the largest number of counts as that to assign to each quality principle, the nine evaluated quality principles were exclusively categorized as *must-be* (see Table 5). This gives all attributes equal impact, which made it impossible to prioritize certain quality principles as desired, despite differences in ratings.

XSL•FO
RenderX

Table 5. Categorization of the answers for the functional and dysfunctional questions related to the nine quality principles, based on the category with the maximum count.

Quality principle	Test group, A (n=191)							Validation group, B (n=191)						
	M[a]	P[b]	A[c]	I[d]	R[e]	Q[f]	Category	M	P	A	I	R	Q	Category
Practicality	127	42	10	7	2	3	M	122	48	4	12	2	3	M
Risk adequacy	127	48	2	9	1	4	M	127	46	0	7	3	8	M
Ethical soundness	120	40	8	19	1	3	M	123	33	7	23	0	5	M
Legal conformity	148	27	2	13	0	1	M	146	20	5	15	1	4	M
Content validity	139	42	1	7	0	2	M	140	38	5	6	0	2	M
Technical adequacy	83	68	20	18	2	0	M	89	59	24	16	1	2	M
Usability	103	49	20	17	0	2	M	105	50	16	15	0	5	M
Resource efficiency	63	40	45	40	1	2	M	69	37	34	40	6	5	M
Transparency	103	43	18	23	3	1	M	89	45	22	27	1	7	M

[a]M: must-be.
[b]P: performance.
[c]A: attractive.
[d]I: indifferent.
[e]R: reverse.
[f]Q: questionable.

For example, for resource efficiency, less than half as many answer pairs were categorized under *must-be* compared with those for content validity (Group A: 63 vs 139 or 45.3%; Group B: 69 vs 140 or 49.3%); nevertheless, both principles were still equally categorized as *must-be*.

If-Then–Based Approach

The situation did not improve when employing the if-then approach; the results were equivalent to those shown in Table 5.

Timko Approach

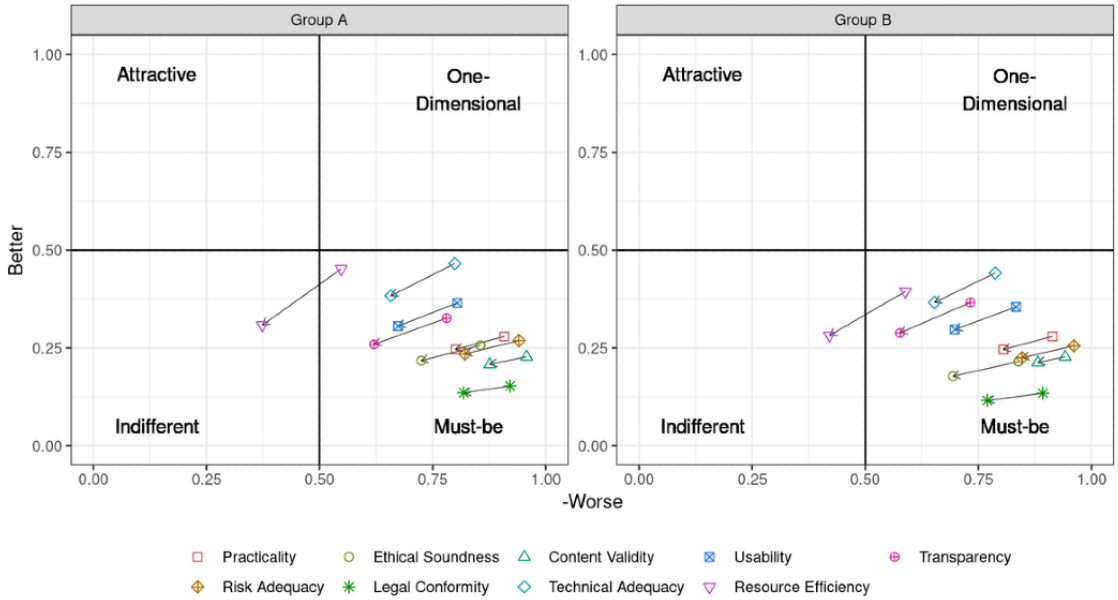
Even using the method proposed by Timko [25], with or without using the average values for perceived importance, the situation only changed marginally, as shown in Table 6 and Figure 6. Visually, the value pairs were still in close vicinity to each other. Without factoring in perceived relevance, all values firmly remained categorized as *must-be*; only when accounting for relevance, one quality principle, specifically resource efficiency, showed a categorization change from *must-be* to *indifferent*. Apart from this principle (which, now being rated *indifferent* is deemed to be of less importance), prioritization of the remaining attributes was elusive, despite apparent (visual and numeric) differences.

Table 6. Better and Worse values without (denoted by a subscripted N) and with factoring in the average value of perceived relevance (or importance, denoted by a subscripted I) for each principle.

Quality principle	Group A					Group B				
	$Better_N$	$Worse_N$	Importance	$Better_I$	$Worse_I$	$Better_N$	$Worse_N$	Importance	$Better_I$	$Worse_I$
Practicality	0.28	–0.91	0.88	0.25	–0.80	0.28	–0.91	0.88	0.25	–0.81
Risk adequacy	0.27	–0.94	0.87	0.23	–0.82	0.26	–0.96	0.88	0.23	–0.85
Ethical soundness	0.26	–0.86	0.85	0.22	–0.72	0.22	–0.84	0.83	0.18	–0.69
Legal conformity	0.15	–0.92	0.89	0.14	–0.82	0.13	–0.89	0.86	0.12	–0.77
Content validity	0.23	–0.96	0.91	0.21	–0.88	0.23	–0.94	0.94	0.21	–0.88
Technical adequacy	0.47	–0.80	0.82	0.38	–0.66	0.44	–0.79	0.83	0.37	–0.65
Usability	0.37	–0.80	0.84	0.31	–0.67	0.35	–0.83	0.84	0.30	–0.70
Resource efficiency	0.45	–0.55	0.68	0.31	–0.37	0.39	–0.59	0.71	0.28	–0.42
Transparency	0.33	–0.78	0.79	0.26	–0.62	0.37	–0.73	0.79	0.29	–0.58

XSL•FO
RenderX

Figure 6. Better and Worse pairings for the training (Group A) and validation (Group B) groups, plotted with and without the average value for perceived importance. The arrows represent the corresponding coordinate shift from the original values to those factoring in the perceived importance for each quality principle.



In-Line-of-Sight Method

Table 7 shows the rankings for both groups based on the *must-be* quadrant, as this is where the attributes predominantly clustered. Angles were calculated in the direction of the *one-dimensional* (*performance*) category.

The distances between Better-Worse pairings for both groups (ie, the distance between the two groups) only differed insignificantly: they always remained below 5% the maximum possible distance within the coordinate square (ie, $0.05\times[(0,0),(-1,1)]=0.05\times\sqrt{2}\approx0.05\times1.14142\approx0.0707$).

Based on the described method, the ranking for the quality principles was identical for both groups, with legal conformity ranked first, followed by content validity, risk adequacy, practicality, ethical soundness, usability, transparency, technical adequacy, and finally, resource efficiency.

Table 7. Ranking the quality principles based on distance to the must-be corner and angle toward the right-most boundary.

Quality principle	Coordinate distance between groups	Group A (test group)				Group B (validation group)			
		Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank	Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank
Practicality	0.00	0.32	51	0.36	4	0.31	52	0.35	4
Risk adequacy	0.03	0.29	53	0.34	3	0.27	56	0.32	3
Ethical soundness	0.05	0.35	38	0.38	5	0.35	30	0.38	5
Legal conformity	0.05	0.23	37	0.26	1	0.26	27	0.28	1
Content validity	0.01	0.24	59	0.29	2	0.24	61	0.29	2
Technical adequacy	0.02	0.51	48	0.55	8	0.50	47	0.54	8
Usability	0.03	0.45	43	0.48	6	0.42	45	0.46	6
Resource efficiency	0.05	0.70	26	0.72	9	0.64	26	0.66	9
Transparency	0.05	0.46	34	0.49	7	0.51	34	0.54	7

Gender Influence

There was only a slight difference in the quality principle–related assessments between male and female participants. As there were too few female participants to prevent outliers from unduly influencing the results to continue evaluating groups A and B separately in this regard, the overall group of all participants was stratified by gender. There were only small differences in prioritization, despite (significant) disparities between both strata regarding the actual placement of the principles in the coordinate system (Figure 7 and Table 8).

XSL•FO
RenderX

Figure 7. Plot of the Better and Worse coordinates per principle stratified by gender.

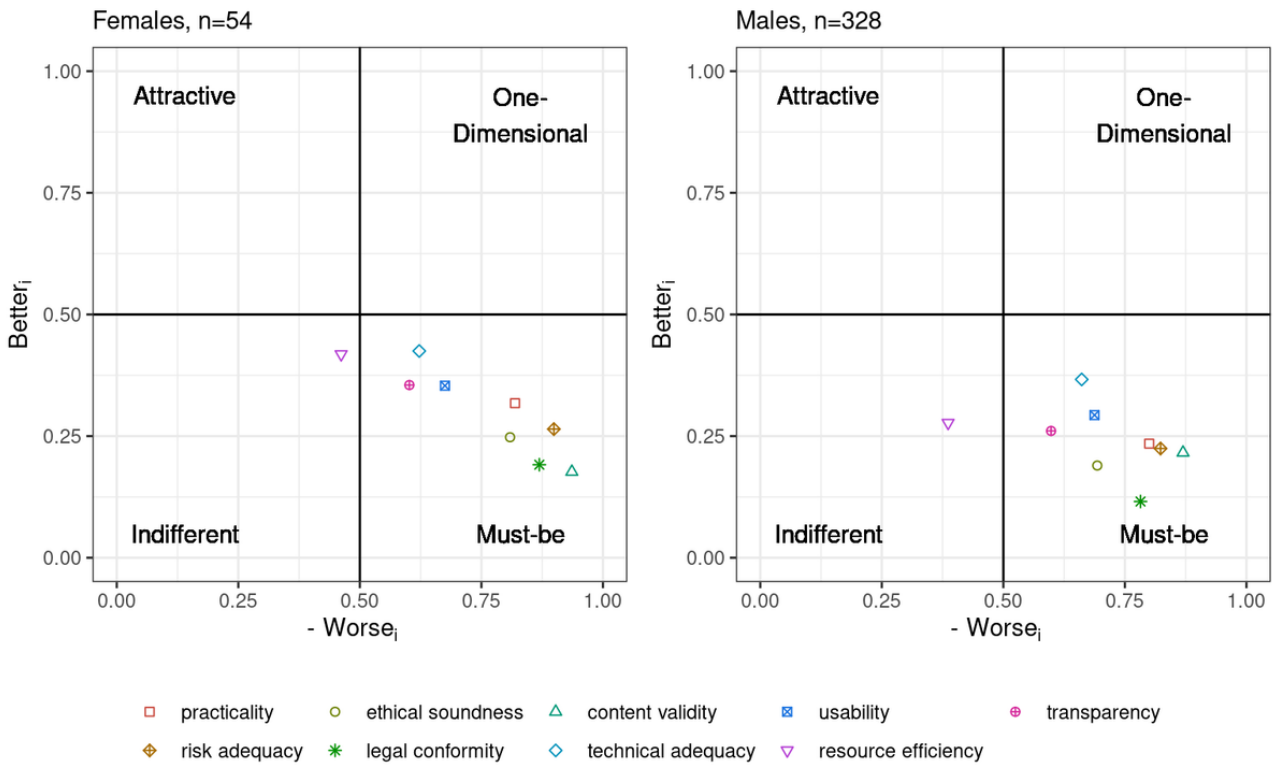


Table 8. Ranking of the quality principles based on the distance of the Better and Worse coordinates to the outermost corner of the must-be quadrant, using the in-line-of-sight method for all participants, stratified by gender.

Quality principle	Coordinate distance between strata	Female participants				Male participants			
		Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank	Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank
Practicality	0.085	0.37	60	0.41	5	0.31	50	0.35	4
Risk adequacy	0.086	0.28	69	0.34	3	0.29	52	0.33	3
Ethical soundness	0.130	0.31	52	0.35	4	0.36	32	0.39	5
Legal conformity	0.115	0.23	56	0.28	2	0.25	28	0.27	1
Content validity	0.077	0.19	70	0.24	1	0.25	59	0.30	2
Technical adequacy	0.070	0.57	48	0.61	8	0.50	47	0.54	8
Usability	0.062	0.48	47	0.52	6	0.43	43	0.46	6
Resource efficiency	0.160	0.68	38	0.71	9	0.67	24	0.69	9
Transparency	0.094	0.53	42	0.57	7	0.48	33	0.51	7

Stratification by Interest in Digitization

There were notable differences in ratings between those with a stated interest in digitization and those who lacked interest in this topic, again considering only the overall group and discarding groups A and B due to the low number of participants in the "little to no interest" stratum (Figure 8). For the latter group, the principles were almost exclusively located in the *indifferent* quadrant, or, in the case of legal conformity, content validity, and risk adequacy, near the border between the *indifferent* and *must-be* quadrants.

Nevertheless, the prioritization remained largely similar with that of the interest-based stratification, with only minor differences (see Table 9).

XSL•FO
RenderX

Figure 8. Plot of the Better and Worse coordinates per principle stratified by interest in the topic.

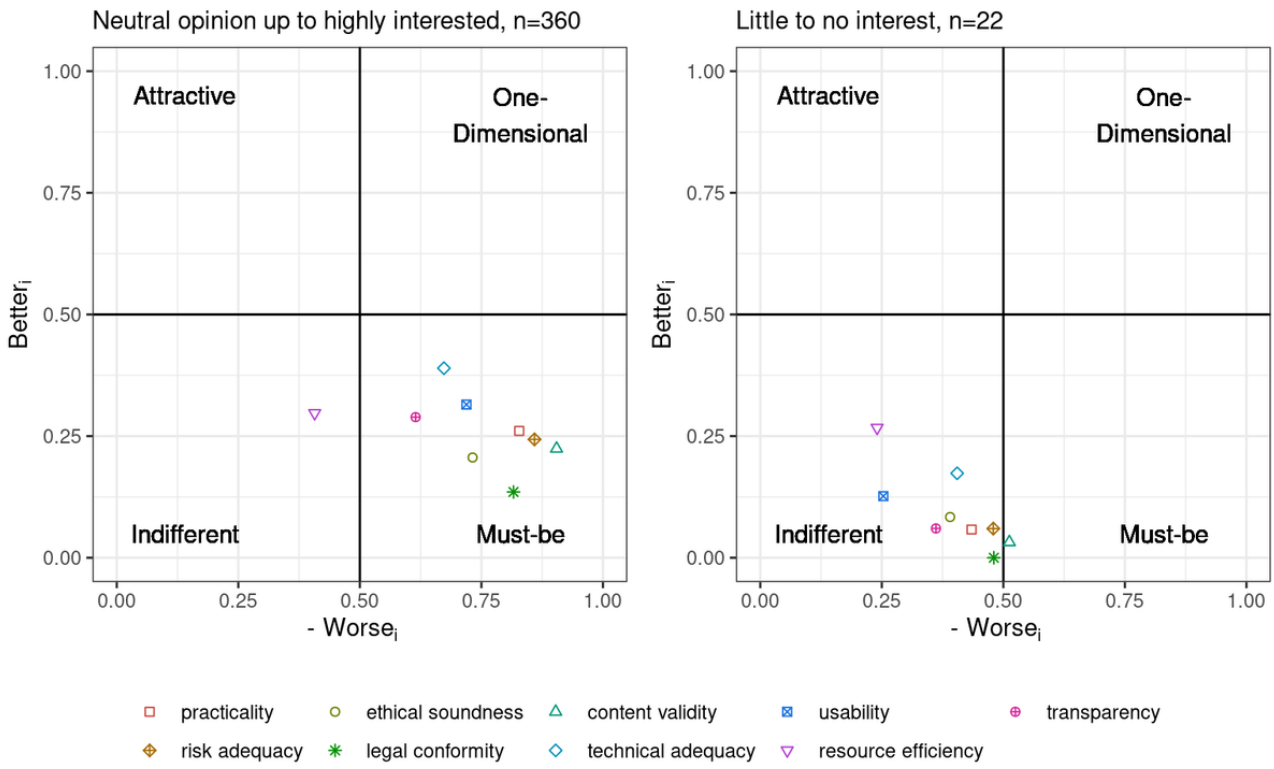


Table 9. Ranking of the quality principles based on the distance of the Better and Worse coordinates to the outermost corner of the must-be quadrant, using the in-line-of-sight method for all participants, stratified by their interest in digitization.

Quality principle	Coordinate distance between strata	Interested participants				Uninterested participants			
		Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank	Distance, *d*	Angle, α	Ranking coefficient, *f*	Rank
Practicality	0.44	0.31	57	0.36	4	0.57	5.9	0.57	4
Risk adequacy	0.42	0.28	60	0.33	3	0.52	6.6	0.53	3
Ethical soundness	0.36	0.34	38	0.37	5	0.61	7.8	0.62	5
Legal conformity	0.36	0.23	36	0.26	1	0.52	0.0	0.52	2
Content validity	0.44	0.24	67	0.30	2	0.49	3.8	0.49	1
Technical adequacy	0.34	0.51	50	0.55	8	0.62	16.3	0.63	6
Usability	0.50	0.42	48	0.46	6	0.76	9.6	0.77	8
Resource efficiency	0.17	0.66	27	0.68	9	0.81	19.4	0.82	9
Transparency	0.34	0.48	37	0.51	7	0.64	5.4	0.65	7

Discussion

Principal Results

As shown in the literature (eg, [23,25,37,38]) as well as our own results, established methods for working with the results of Kano surveys are well-suited to determining generic user perceptions of product attributes of a health app, such as the quality principles that the participants of our survey were confronted with.

Nevertheless, when using Kano's original approach, or even the more promising approach proposed by Timko [25] (with or without inclusion of the perceived relevance of the principles), in our case, the nine attributes remained firmly tethered to the *must-be* category (see Figure 6), with only resource efficiency crossing into the *indifferent* realm once perceived importance was included in the calculation. However, there were no *one-dimensional* or even *attractive* attributes. Solely based on established evaluation methods for Kano surveys, we therefore

XSL•FO
RenderX

fell short of obtaining the desired ranking to be used for potentially prioritizing the health app quality principles.

Simply applying the Kano method and its categorizations to the quality principles initially did not allow for prioritization, which confirmed the previously noted similarity of the ratings [18,19], with again only resource efficiency standing out. As reported previously [18,19], the discrepancy between this quality principle and the other eight principles supports the assumption that resource efficiency likely only plays a minor role in today's mostly very powerful devices, since health-related apps in particular presumably place little demand on the devices.

To counteract this lack of differentiation between the principles, we then developed the so-called "in-line-of-sight" method, which, based on the numeric values representing satisfaction as well as dissatisfaction with the respective attribute or quality principle, determines a ranking coefficient while also accounting for different points of view (depending on the purpose of the desired prioritization). This method should also be flexible enough to be adapted to different circumstances depending on the use case and user ratings provided.

In our exemplary evaluation for the ranking from the *must-be* perspective, we chose a rather conservative approach, factoring in an angle that leads to lower Worse values being preferred, while accepting that by choosing this approach, the values for Better will also decrease.

This corresponds to the definition of the *must-be* category: a lack of the respective quality principle would be perceived more strongly than the positive effect that would be achieved if the characteristics consistent with the quality principle were present. When changing the perspective to another quadrant, similar considerations need to be applied, with calculations being adapted accordingly. For example, when changing the perspective to the *attractive* quadrant, it would be more useful to aim at a higher priority of Better values, as this better represents the definition of this category.

Kano Survey Interpretation: Potential for Linguistic Inconsistencies?

Although the Kano model is popular and is often used in a wide variety of contexts, linguistic inaccuracies in its application have arisen over the years, which in some publications have led to difficulties in its correct application or to supposed inconsistencies ([29], citing [25]). The problem originates from an inaccurate translation of Kano's key concept transliterated as "atarimae," which has been translated as *must-be* in many English-language publications. *Must-be* seems to have first been used in the early 1990s by Shoji Shiba when presenting the Kano model to English-speaking audiences [29]. However, apparently, the meaning of "atarimae" would be better represented by the terms "natural," "obvious," "expected," "ordinary," or "normal." This change should be applied to the category name *must-be* as well as the corresponding customer response, which is often given as "It must be that way," but, as noted by Horton and Goers [29], should rather be represented by translations along the lines of the aforementioned suggestions.

When Kano surveys are translated into other languages, this inaccuracy may be passed on to a varying degree, potentially further complicating the situation. In our (German language) questionnaire, however, we already included the wording representing "I take this for granted" (German: "Setze ich voraus") as an answer option for the participants instead of *must-be*, thus more closely following Kano's original idea. To stay in line with most of the literature, we nevertheless decided to stick to the *must-be* term, although this aspect needs to be kept in mind. This change in interpretation may also provide an explanation for the results we obtained for the nine quality principles, with all of them being located in the *must-be* category.

In contrast to common usage scenarios for Kano surveys that aim at selecting attributes one should further investigate, we applied the model to a set of attributes, namely our quality principles, that had already been painstakingly compiled [16,17] (among others based on various norms (eg, [3,39-42]), as well as the literature (eg, [5,6,43,44]). This may provide an additional explanation for why, in the survey presented here, all quality principles were rated as *must-be*, or following the adapted interpretation, as "obvious" or "something to be taken for granted." That is, the quality principles simply followed obvious requirements that were mentioned as essential in the aforementioned sources, and that one would expect users to be able to rate objectively (at least to a certain degree); they were, however, not selected in order to trigger enthusiasm. Their placement in the *must-be* quadrant is therefore easily explained, and the sole exception for resource efficiency being placed in the *indifferent* category may possibly be due to the fact that today's mobile devices are commonly equipped with sufficient computing power—at least for physicians, who often probably have access to rather high-end devices—so that resources are not a factor that warrants considerable attention.

Selection of the Evaluation Method Used as a Basis of this Work

In addition to the linguistic aspects, there is no clear verdict about the methodology one should apply foremost when evaluating Kano model–based surveys. While there is a large variety of methods to choose from, based on various theoretical concepts, the discussion is still open as to which of them is most appropriate (in general or for a specific use case) and has the greatest validity. Although there are various empirical evaluations of different approaches in the context of Kano surveys that are described in the literature (eg, [37,38,45-47]), determining which of these particular approaches is best seems to be near impossible.

As stated by Mikulić and Prebezac [23], the validity and reliability of the various approaches cannot be determined with certainty: there is simply no known comparison that can be taken as the ground truth.

Which method is chosen is therefore rather often a matter of whether (1) the theoretical justification of the respective approach appears valid, (2) the increase in information when applying the respective approach actually contributes to the solution of the problem, and (3) which (recognizable) technical strengths and weaknesses the approach has.

XSL•FO
RenderX

For the purposes of this paper, Timko's approach (first introduced in [25]) was therefore chosen as a foundation, as it is easy to understand and also easily allows for integration of the self-stated relevance of the attributes to be evaluated. Additionally, compared to Kano's initial idea, where, essentially, all 25 possible answer combinations are directly mapped to only 6 possible categories, one may feel the need for a more differentiated, continuous method of analyzing the data to better assess how different attributes are similar or dissimilar, and our enhanced approach follows this line of thought.

Limitations

Selection of the Quality Principles Employed in This Study

New information technologies, including online information or specific (mobile) apps, place additional demands on those employing them, especially in professional health care contexts. Professionals employing such technologies need to ensure that they are safe and pose no harm to those in their care. Regulatory oversight as well as evidence-based literature are often found lacking [48]. Economic questions such as the paucity of information related to cost-effectiveness or cost-utility [49,50], or even aspects related to reimbursement [51,52] may also play a role in whether or not the technologies are actually adopted in everyday practice.

Without at least a basic understanding of the relevant quality aspects (and how to apply them), or uncertainties regarding their safety and security, acceptance may suffer, which may also limit the potential of these technologies [48,53]. However, there is no general consensus, even among experts, as to what exactly constitutes "quality" in this context and how it can be assessed for specific scenarios (eg, to rate health-related apps) [54].

To identify items of relevance, such as for inclusion in various tools [4-8] meant to aid in assessing such technologies that are to be provided to the respective target groups (eg, physicians or other health care personnel), it is important to identify certain key aspects in the hope that these fulfill the information needs and information-seeking behaviors of users [55]. Many authors use rather detailed approaches and criteria to enable this information-seeking and more easily assess the quality of health-related apps, and they often target specific (professional) user groups [54].

For this purpose, in close collaboration with various stakeholders (eg, experts convened on behalf of eHealth Suisse), the nine quality principles used here were compiled [16,17] and evaluated [18,19]. In this context, we were able to show that, despite its broad scope and lack of details, and being almost unanimously regarded as (highly) relevant by the participants of both previous studies, the predefined set of quality principles was still well-suited to provide the respective participants with pointers to aspects relevant for determining an app's quality and fine-tuning their usage decisions. After having been sensitized to the topic of quality principles, and having applied these principles to exemplary app descriptions, the participants of both previous studies were able to make a much more differentiated assessment of the app descriptions that were provided, and were much more critical in their decision on whether or not to potentially use the corresponding app.

Survey Design

Although we had initially considered an additional qualitative approach, specifically to ask the participants to directly rank the principles as they saw fit, a major reason that made us abandon this course of action was that the data presented here were part of a larger project (as mentioned above, the first part of the analysis of the acquired data is already published [28]), and it was decided by the team that an additional (sorting) questionnaire would be too much of a burden for those participating in the survey. Of the two alternatives for designing the part of the survey presented here (ie, continuing to rely on the Kano model or using the qualitative sorting approach), the choice ultimately fell on Kano. This was based on our hope to be able to use the data obtained for implicit assessment instead of running the risk that the previously established, highly similar assessments of the principles would make it difficult for the participants to determine a specific order. Because we did not initially know how many people would participate, we were concerned that it would be difficult to determine an overall ranking for the nine principles if too few people participated and we only relied on the explicitly stated rankings. It was hoped that based on Kano's methods, using the provided answers and ensuing categorizations, we would be able to at least determine a rough prioritization for the overall group of participants, in our case, by giving principles in the *must-be* or *one-dimensional* categories precedence over those in the *attractive* or *indifferent* categories.

Study Participants

Despite having contacted a relatively large number of potential participants, with only 4.02% (382/9503) of those who were initially invited actually completing the survey, the response rate was low. Based on this response rate and demographic factors, the results, specifically those related to any rankings of attributes presented here, may not be fully representative of physicians overall or even those specializing in orthopedic or trauma surgery.

One of the possibly most relevant demographic factors for which one might potentially expect an impact on the assessments is the gender of the participants. Overall, the gender distribution of the participants roughly corresponded to the ratio expected in orthopedics. In our survey, 85.9% (328/382) of the participants were male and 14.1% (54/382) were female. Thus, there were only slightly fewer women than would have been expected in the field of orthopedics and trauma surgery, according to data provided by the Bundesärztekammer, with 17.63% (3611/20,477), as of December 31, 2020, of those in the fields of orthopedics or orthopedics and trauma surgery being women [56].

However, gender seems to only have exerted a limited influence on prioritization, which is in line with our previous work [18,19], where there were also only minor differences in the quality principle–related assessments between male and female participants. Differences were particularly pronounced for resource efficiency and ethical soundness (see the column

XSL•FO
RenderX

describing the coordinate distance between both strata in Table 8, as well as Figure 7 for the actual coordinates). The former was placed near the (neutral) center for female participants, whereas for male participants, it was clearly placed in the *indifferent* quadrant. Content validity and usability (along with transparency for female participants) were somewhat closer to the *one-dimensional* quadrant than the other principles in both strata. In case of the female participants, the point cloud was also shifted more toward the *one-dimensional* quadrant compared with that of their male peers, and the coordinates were less scattered overall (Figure 7).

Regarding the ranking of the principles, for the female participants (n=54), content validity ranked first and legal conformity ranked second (Table 8). For the male (n=328) participants, this order was reversed. The same was true for ethical soundness and practicality. Apart from resource efficiency, all quality principles were found in the *must-be* quadrant (Figure 7).

Nevertheless, the prioritization was roughly similar for the two demographic groups: for the female participants (n=54), content validity ranked first and legal conformity was placed second (Table 8), whereas this order was reversed for the male (n=328) participants. The same was true for ethical soundness and practicality. Apart from resource efficiency, all quality principles were found in the *must-be* quadrant (Figure 7).

Considering interest in digitization (Figure 8 and Table 9), digitally affine participants (aggregated data for "neutral," "interested," or "highly interested"; n=360) were considerably overrepresented due to the chosen survey method. Participants with little interest in the topic, or those lacking access to the techniques used, responded much less frequently than those showing more enthusiasm toward digitization, thus potentially biasing the results as well. However, for the limited number of participants (n=22) who cared only little about digitization (values aggregated for being "less interested" or "not interested" in the topic), but nevertheless participated, it was primarily the placement of the points representing the quality principles in the coordinate system that differed strikingly from the other participants (Figure 8). There was also a striking difference in the placement of the principles within the coordinate system, which is probably not solely attributable to the imbalance between the sizes of the two groups. Disinterested participants rated the principles as *indifferent*, or, in the case of legal conformity, content validity, and risk adequacy, near the border between the *indifferent* and *must-be* quadrants (Figure 8). Nevertheless, rankings remained largely similar independent of digital affinity. For those stating a more or less pronounced interest into digitalization, the order of practicality and risk adequacy was reversed compared with that of the participants with little to no interest. Among disinterested participants, there were also small deviations in the rank for legal conformity and content validity (reverse rank 1 and 2, respectively) as well as technical adequacy and usability (rank 6 and 8, respectively; see Table 9). Legal conformity, content validity, and risk adequacy occupied the top ranks among participants with or without interest, but the order for content validity and legal conformity differed. The lower ranks were occupied by usability, transparency, technical adequacy, and resource efficiency, albeit with a somewhat differing order.

The difference in locations of the principles in the coordinate system (Figure 8), but not in the prioritizations obtained for the two groups (Table 9), lends support to the feasibility of applying our method to quality principles in the mHealth app domain, and supports the need for better education of (potential) users of mHealth apps. Although medical professionals such as our participants are—or at any rate should be—aware of the need for quality (as demanded by professional ethics) for all tools they apply in care contexts, it seems as though for those lacking interest in digitization, this mental transfer apparently does not work for the uninterested participants, as shown by their indifferent ratings. Educational campaigns such as those by professional societies that emphasize the need for quality not only in conventional care but also in the digital domain, including mHealth apps, may help to raise awareness in this regard even for those who are not (yet) familiar or comfortable with the use of such technologies in their daily work.

Altogether, an additional, hopefully larger-scale, study should be implemented to obtain more conclusive data for these as well as other demographic strata, such as by recruiting additional participants with the aid of other professional organizations or by including additional target groups such as patient organizations, universities providing medical education, and others.

Implementation

We believe to have found a methodology that is well-adapted to the demands of finding a prioritization of app quality principles in the case of very similar categorizations, clustered in either of the four categories of *must-be*, *one-dimensional*, *attractive*, or *indifferent* obtained using a Kano questionnaire.

Of course, our method needs further validation, and, depending on the scenario in which it is applied, it might be helpful to adapt the strategy of how the angles (or their direction) are calculated. This may depend on multiple factors. For example, when considering ratings based on *must-be*, it seems sensible to always perform sorting based on the distance to the *one-dimensional* rather than to the *indifferent* quadrant, as an *indifferent* opinion, per se, does not elicit identification with the product (or its attributes).

However, if one switches perspective to the *one-dimensional* category, it may well depend on the type of product, its application areas, as well as its target user group, along with the attributes actually being evaluated if it makes more sense to calculate the angles used in determining the sorting against *must-be* or *attractive*. For products targeting professionals, it might, for example, make more sense to sort the quality principles depending on their closeness to *must-be*, whereas for marketing purposes, *attractive* qualities may be more promising. Again, for the *attractive* corner, similar arguments as for the approach taken for *must-be* apply, with the angle toward *one-dimensional* rather than *indifferent*, which likely makes more sense in most scenarios.

If attributes were clustered in the *indifferent* corner, the question of the direction to base any attribute sorting on is again more

XSL•FO
RenderX

open (ie, toward either *attractive* or *must-be*). The decision may also depend somewhat on the purpose, design, and area of use of the product under consideration; in the case of a professional product, it may potentially make more sense to build the ranking based on *must-be* as a reference, since *attractiveness* does not necessarily reflect professional quality.

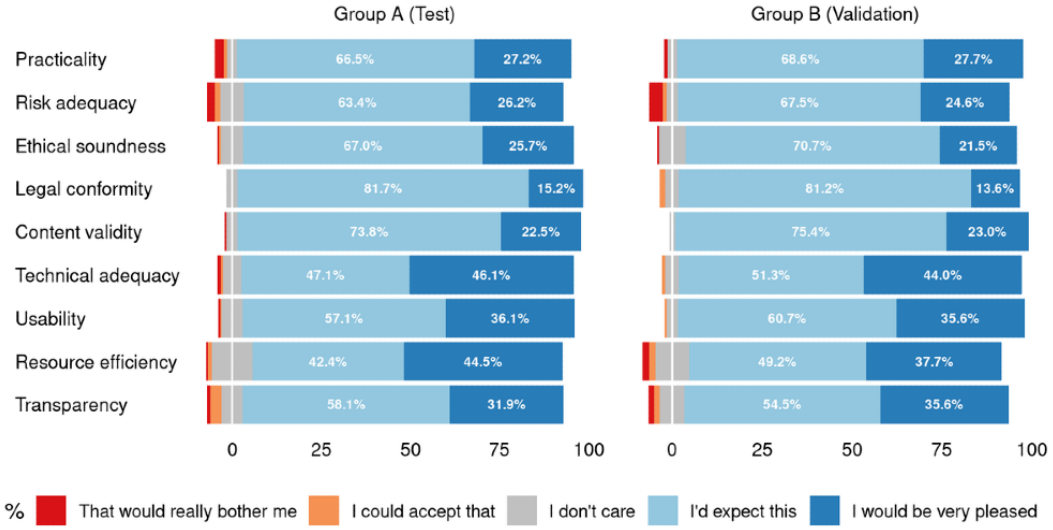
Outlook and Comparison With Previous Work

Further proof of the validity of the method and its transferability to other interest groups, quality attributes, or application scenarios is still pending. Future work will particularly have to address further validation of the method with regard to the evaluation involving other user groups (eg, patients, caregivers) or to the application for prioritization of other attributes, whether for use in medical or general apps, or for the evaluation of other attribute lists outside the app domain.

However, especially with regard to the determined ranking of the quality criteria we chose for this evaluation, we believe that a comparison of the perception of relevance between the results of the previous studies (eg, [19], where participants were working in a different medical field) and those shown here is a strong indicator that the results are likely transferable. Similar to the current work, participants of previous studies had also been asked to provide their opinion regarding the relevance of the nine quality principles, and the participating physicians rated the relevance of the quality principles similar to the current group of participants (see Figure 9), with only minor (and statistically negligible) differences between the previous study [19] and the data obtained from the participants of this study. Table S4 in Multimedia Appendix 1 shows the overall relevance ratings and *P* values for the comparison between the two studies. However, for the sake of streamlining the comparison between studies, the respective test and validation samples, as they were used in both studies, were aggregated. Similarly, to stay in line with Albrecht et al [19], the answer options for "very important" and "important" were summarized using the term "important," while those for "less important" and "unimportant" were aggregated as "not important."

As shown in Figure 9 and Table S4 in Multimedia Appendix 1, there are notable similarities between both studies: the proportion of participants that rated *resource efficiency* as important was decidedly lower (current study: 260/382, 68.1% participants; previous study: 270/441, 61.6%) than it was for all other quality principles, where the perceived importance was in the range of 84%-98%, again for both studies.

Figure 9. Relevance ratings for the nine quality principles: comparison between this survey and previously published work [19]. See Table S4 in Multimedia Appendix 1 for the corresponding *P* values of this comparison. DGIM: German Association for Internal Medicine, German: "Deutsche Gesellschaft für Innere Medizin e.V.".



Conclusions

The agreement with respect to perceived relevance between both studies, as shown above, leads to the following conclusions.

For both previous studies [18,19], there was no clear pathway for prioritization of the principles should the need arise, apart from resource efficiency consistently being the least popular quality principle, with ensuing lesser relevance. However, in today's medical world, time is a valuable commodity, and in fact, a lack of time or too much effort being required to adequately assess all relevant aspects is often mentioned as a barrier both to accessing information [55] as well as to employing apps in specific situations (eg, for consultations [57]). Although health apps may initially give the impression of being able to save time and reduce effort, professional ethics (eg, [58,59]) demand that those working in medical professions must ensure that any (digital) tools they use are up to the expected professional standards. In the digital world, even aided by various tools meant to aid in the process, health care professionals often remain unsure of which factors they need to consider in this context, especially if the tools require extensive effort. This may possibly contribute to the many—real or perceived—barriers toward successfully using apps in care settings or for health-related purposes in general.

Of course, an all-encompassing, unaided, and professionally conducted evaluation of apps will neither be possible nor practical in most scenarios, largely due to a lack of technical expertise. However, physicians and other health care professionals should at least be enabled to assess available

XSL•FO
RenderX

information in the context of their work, such as based on a set of questions [19] that address basic quality principles. Even for such limited lists, being able to determine a ranking of the questions or quality principles seems sensible for assessing highly available information with priority; if the initially evaluated factors already lead to a rejection, the remaining factors can justifiably be disregarded, thus saving the time that a full, structured assessment based on such questions covering all available information sources (eg, from the app store, on manufacturer websites, and other sources) would take. For longer lists of quality principles or rating criteria applied to mHealth apps, the benefits of being able to determine a sensible and context-adapted prioritization, based on feedback obtained from the respective peer group, may be even greater, counteracting or at least somewhat alleviating arguments that many of the available rating tools or quality principles are—due to the large number of details they cover—too cumbersome for real-word applications outside of academic evaluations [22].

In contrast to other approaches based on the Kano method (eg, [23,25]) that predominantly strive for categorization of the attributes being evaluated, the methodology presented here may provide an interesting option that additionally allows for the prioritization of quality principles in cases of largely similar categorization results or initial user perceptions. This may aid in giving precedence to the most relevant (prioritized) principles, deferring those with lesser priority. To what extent the method will be applicable beyond the usage scenario described here will require more extensive investigations.

However, it also remains an open question as to how one could deal with cases where for a larger number of attributes, there are multiple close clusters of attributes found in different quadrants. One possible solution to this might be to sort attributes in each cluster as described above, and to then perform a prioritization of the clusters themselves (with attributes in the *attractive* quadrant probably being the most relevant) in order to arrive at a full ranking of all attributes to be considered.

Nevertheless, the proposed prioritization may provide a means for professional organizations that want to give their members a recommendation as to which quality principles should be applied with priority in digital domains, independent of whether this is done for the generic set of app-related quality principles or principles that are more subject-specific (eg, for use in a particular medical specialty or for a specific user group).

Acknowledgments

The authors would like to thank the DGOU and the BVOU for the logistical support of the survey. Special thanks also go to Prof Bernhard Breil for the valuable discourse.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables.
[PDF File (Adobe PDF File), 1735 KB - mhealth_v10i1e26563_app1.pdf]

References

1. Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
2. Hillebrand U, von Jan U, Albrecht UV. Concepts for quality assurance of health related apps. Stud Health Technol Inform 2016;226:209-212. [Medline: 27350506]
3. British Standards Institute Staff. PAS 277:2015 – Health and wellness apps – Quality criteria across the life cycle – Code of practice. London, UK: British Standards Institute; Apr 30, 2015.
4. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
5. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
6. Nouri R, Kalhori SRN, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Inform Assoc 2018 Aug 01;25(8):1089-1098 [FREE Full text] [doi: 10.1093/jamia/ocy050] [Medline: 29788283]
7. O'Rourke T, Pryss R, Schlee W, Probst T. Development of a multidimensional app-quality assessment tool for health-related apps (AQUA). Digit Psych 2020 Oct 27;1(2):13-23 [FREE Full text] [doi: 10.24989/dp.v1i2.1816]
8. Llorens-Vernet P, Miró J. The Mobile App Development and Assessment Guide (MAG): Delphi-based validity study. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e17760 [FREE Full text] [doi: 10.2196/17760] [Medline: 32735226]

XSL•FO
RenderX

9. Albrecht UV, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: analysis. JMIR Mhealth Uhealth 2018 Apr 25;6(4):e10394 [FREE Full text] [doi: 10.2196/10394] [Medline: 29695374]
10. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
11. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
12. Radovic A, Badawy SM. Technology use for adolescent health and wellness. Pediatrics 2020 May;145(Suppl 2):S186-S194. [doi: 10.1542/peds.2019-2056G] [Medline: 32358210]
13. Rigamonti L, Albrecht UV, Lutter C, Tempel M, Wolfarth B, Back DA, Working Group Digitalisation. Potentials of digitalization in sports medicine: a narrative review. Curr Sports Med Rep 2020 Apr;19(4):157-163. [doi: 10.1249/JSR.0000000000000704] [Medline: 32282462]
14. Albrecht UV, Pramann O, von Jan U. Synopsis for health apps: transparency for trust and decision making. In: Househ M, Borycki E, Kushniruk A, editors. Social media and mobile technologies for healthcare. Hershey, PA: IGI Global; 2014:94-108.
15. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
16. Albrecht UV. Catalogue de critères uniformes pour l'autodéclaration de la qualité des applications de santé. Bern, CH: ehealth Suisse; 2019.
17. Albrecht UV. Einheitlicher Kriterienkatalog zur Selbstdeklaration der Qualität von Gesundheits-Apps. Bern, CH: eHealth Suisse; 2019.
18. Albrecht U, Malinka C, Long S, Raupach T, Hasenfuß G, von Jan U. Quality principles of app description texts and their significance in deciding to use health apps as assessed by medical students: survey study. JMIR Mhealth Uhealth 2019 Feb 27;7(2):e13375 [FREE Full text] [doi: 10.2196/13375] [Medline: 30810534]
19. Albrecht UV, Framke T, von Jan U. Quality awareness and its influence on the evaluation of app meta-information by physicians: validation study. JMIR Mhealth Uhealth 2019 Nov 18;7(11):e16442 [FREE Full text] [doi: 10.2196/16442] [Medline: 31738179]
20. Kano N, Seraku N, Takahashi F, Tsuji SI. Attractive quality and must-be quality. J Japan Soc Qual Control 1984 Apr;14(2):147-156.
21. Sauerwein E. Das Kano-Modell der Kundenzufriedenheit. In: Das Kano-Modell der Kundenzufriedenheit: Reliabilität und Validität einer Methode zur Klassifizierung von Produkteigenschaften. Wiesbaden, Germany: Deutscher Universitätsverlag; 2000:27-55.
22. Guo Y, Bian J, Leavitt T, Vincent HK, Vander Zalm L, Teurlings TL, et al. Assessing the quality of mobile exercise apps based on the American College of Sports Medicine guidelines: a reliable and valid scoring instrument. J Med Internet Res 2017 Mar 07;19(3):e67 [FREE Full text] [doi: 10.2196/jmir.6976] [Medline: 28270378]
23. Mikulić J, Prebežac D. A critical review of techniques for classifying quality attributes in the Kano model. Manag Serv Qual 2011 Jan 25;21(1):46-66. [doi: 10.1108/09604521111100243]
24. Kano N, Seraku N, Takahashi F, Tsuji S. Chapter 10: Attractive quality and must-be quality. In: Hromi J, editor. The best on quality: targets, improvements, systems. Volume 7. Milwaukee, WI: Quality Press; 1996:165-186.
25. Berger C, Blauth R, Boger D, Bolster C, Burchill G, DuMouchel W, et al. Kano's methods for understanding customer-defined quality. Center Qual Manag J 1993;2(4):3-36 [FREE Full text]
26. Leiner DJ. SoSci Survey – the Solution for Professional Online Questionnaires. Version 3.2.000. URL: https://www.soscisurvey.de/en/index [accessed 2021-12-29]
27. Deutscher Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz - DVG). Bundesgesetzblatt 2019 Dec 09;I(49):2562-2584 [FREE Full text]
28. Dittrich F, Albrecht UV, von Jan U, Malinka C, Ansorg J, Jung J, et al. The Digital Healthcare Act - a turning point in the German digitisation strategy? Z Orthop Unfall 2021 Jun;159(3):259-265. [doi: 10.1055/a-1141-4274] [Medline: 32365397]
29. Horton G, Goers J. A revised Kano model and its application in product feature discovery. ResearchGate. 2019. URL: https://www.researchgate.net/publication/332304132_A_Revised_Kano_Model_and_its_Application_in_Product_Feature_Discovery [accessed 2021-12-29]
30. Ernst K. Kategorisierung von Wohnungsmerkmalen in Bezug auf Mieteranforderungen (Bachelor's Thesis). Mittweida, Germany: Hochschule Mittweida; 2015. URL: https://monami.hs-mittweida.de/frontdoor/deliver/index/docId/7019/file/Bachelorarbeit_K.Ernst.pdf [accessed 2021-12-29]
31. Sari Y. The Selection of Classification Methods in Kano Model for Service Industry. 2015 Presented at: 12th Ubaya International Annual Symposium on Management: Sustainability as a Strategic Orientation in Global Economy Recovery Phase; 2015; Surabaya, Indonesia p. 1119-1129.

XSL•FO
RenderX

32. Mkpojiogu EOC, Hashim NL. Understanding the relationship between Kano model's customer satisfaction scores and self-stated requirements importance. Springerplus 2016;5:197 [FREE Full text] [doi: 10.1186/s40064-016-1860-y] [Medline: 27026893]
33. Matzler K, Hinterhuber HH, Bailom F, Sauerwein E. How to delight your customers. J Product Brand Mgt 1996 Apr;5(2):6-18. [doi: 10.1108/10610429610119469]
34. Wickham H, François R, Henry L, Müller K. Dplyr: A grammar of data manipulation. The Comprehensive R Archive Network (CRAN). 2021. URL: https://CRAN.R-project.org/package=dplyr [accessed 2021-12-29]
35. Wickham H, Chang W, Henry L, Pedersen T, Takahashi K, Wilke C, et al. ggplot2: Create elegant data visualisations using the grammar of graphics. The Comprehensive R Archive Network (CRAN). 2021. URL: https://CRAN.R-project.org/package=ggplot2 [accessed 2021-12-29]
36. Heinzen E, Sinnwell J, Atkinson E, Gunderson T, Dougherty G. Arsenal: An arsenal of R functions for large-scale statistical summaries. The Comprehensive R Archive Network (CRAN). 2021. URL: https://CRAN.R-project.org/package=arsenal [accessed 2021-12-29]
37. Bartikowski B, Llosa S. Customer satisfaction measurement: comparing four methods of attribute categorisations. Service Ind J 2004 Jul;24(4):67-82. [doi: 10.1080/0264206042000275190]
38. Löfgren M, Witell L. Two decades of using Kano's theory of attractive quality: a literature review. Qual Manag J 2008;15(1):59-75. [doi: 10.1080/10686967.2008.11918056]
39. ISO/IEC 9126-1 – Software engineering – Product quality – Part 1: Quality model. International Organization for Standardization. 2001. URL: https://www.iso.org/standard/22749.html [accessed 2021-12-30]
40. ISO/IEC 25010 – Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models. International Organization for Standardization. 2011. URL: https://www.iso.org/standard/35733.html [accessed 2021-12-30]
41. ISO/IEC 25023 – Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Measurement of system and software product quality. International Organization for Standardization. 2016. URL: https://www.iso.org/standard/35747.html [accessed 2021-12-30]
42. ISO/IEC 82304-1– Health software – Part 1: General requirements for product safety. International Organization for Standardization. 2016. URL: https://www.iso.org/standard/59543.html [accessed 2021-12-30]
43. Albrecht UV, editor. Chances and Risks of Mobile Health Apps (CHARISMHA). Abridged Version. Hannover, Germany: Medizinische Hochschule Hannover; 2016.
44. Yasini M, Beranger J, Desmarais P, Perez L, Marchand G. mHealth quality: a process to seal the qualified mobile health apps. Stud Health Technol Inform 2016;228:205-209. [Medline: 27577372]
45. Fuchs M, Weiermair K. New perspectives of satisfaction research in tourism destinations. Tourism Review 2003 Mar;58(3):6-14 [FREE Full text] [doi: 10.1108/eb058411]
46. Matzler K, Sauerwein E. The factor structure of customer satisfaction. Int J of Service Industry Mgmt 2002 Oct;13(4):314-332. [doi: 10.1108/09564230210445078]
47. Chen L. A novel approach to regression analysis for the classification of quality attributes in the Kano model: an empirical test in the food and beverage industry. Omega 2012 Oct;40(5):651-659. [doi: 10.1016/j.omega.2011.12.004]
48. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
49. Badawy SM, Kuhns LM. Economic evaluation of text-messaging and smartphone-based Interventions to improve medication adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
50. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
51. Esser M, Boreham A, Ring C, Schreier J. PNS100 The New Reimbursement Route for Digital Health Applications (DIGA) in Germany: critical appraisal and first evaluation of the possible effect on the German healthcare system. Value in Health 2020 Dec;23:S658-S659. [doi: 10.1016/j.jval.2020.08.1544]
52. Degelsegger-Márquez A, Panteli D. Towards reimbursement of mobile digital health applications? An international comparison of practices. Eur J Public Health 2020 Sep;30(Suppl 5):v486. [doi: 10.1093/eurpub/ckaa165.1327]
53. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
54. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. Disabil Rehabil Assist Technol 2021 Oct;16(7):712-721. [doi: 10.1080/17483107.2019.1701103] [Medline: 31910687]
55. van der Keylen P, Tomandl J, Wollmann K, Möhler R, Sofroniou M, Maun A, et al. The online health information needs of family physicians: systematic review of qualitative and quantitative studies. J Med Internet Res 2020 Dec 30;22(12):e18816 [FREE Full text] [doi: 10.2196/18816] [Medline: 33377874]

XSL•FO
RenderX

56. German Medical Association. Ärztestatistik zum 31. Dezember 2020. Berlin, Germany: Bundesärztekammer; 2021. URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Statistik_2020/2020-Statistik.pdf [accessed 2021-12-29]
57. Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among Australian general practitioners: survey study. JMIR Mhealth Uhealth 2019 Jun 03;7(6):e13199 [FREE Full text] [doi: 10.2196/13199] [Medline: 31199343]
58. Chaet D, Clearfield R, Sabin JE, Skimming K, Council on Ethical Judicial Affairs American Medical Association. Ethical practice in Telehealth and Telemedicine. J Gen Intern Med 2017 Oct;32(10):1136-1140 [FREE Full text] [doi: 10.1007/s11606-017-4082-2] [Medline: 28653233]
59. IMIA Code of Ethics for Health Information Professionals. 2016. URL: https://imia-medinfo.org/wp/imia-code-of-ethics/ [accessed 2021-12-29]

Abbreviations

BVOU: Berufsverband für Orthopädie und Unfallchirurgie (Professional Association for Orthopedics and Trauma Surgery)
DGOU: Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (German Society for Orthopedics and Trauma Surgery)
DVG: German Digital Healthcare Act
mHealth: mobile health

Edited by L Buis; submitted 18.12.20; peer-reviewed by P van der Keylen, S Badawy, E Omaki; comments to author 15.04.21; revised version received 10.06.21; accepted 05.11.21; published 11.01.22.

Please cite as:
Malinka C, von Jan U, Albrecht UV
Prioritization of Quality Principles for Health Apps Using the Kano Model: Survey Study
JMIR Mhealth Uhealth 2022;10(1):e26563
URL: https://mhealth.jmir.org/2022/1/e26563
doi:10.2196/26563
PMID:35014965

©Christin Malinka, Ute von Jan, Urs-Vito Albrecht. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 11.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Efficacy, Effectiveness, and Quality of Resilience-Building Mobile Health Apps for Military, Veteran, and Public Safety Personnel Populations: Scoping Literature Review and App Evaluation

Melissa Voth[1,2], BEd; Shannon Chisholm[2], BSc; Hannah Sollid[2], BSc; Chelsea Jones[1,3,4], PhD; Lorraine Smith-MacDonald[1,2], PhD; Suzette Brémault-Phillips[1,2], PhD

[1]Heroes in Mind, Advocacy and Research Consortium, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[2]Department of Occupational Therapy, Faculty of Rehabilitation, University of Alberta, Edmonton, AB, Canada

[3]Leiden University Medical Centre, Leiden University, Leiden, Netherlands

[4]Operational Stress Injury Clinic, Alberta Health Services, Edmonton, AB, Canada

Corresponding Author:
Chelsea Jones, PhD
Heroes in Mind, Advocacy and Research Consortium
Faculty of Rehabilitation Medicine
University of Alberta
1-94 Corbett Hall
8205 - 114 Street Edmonton
Edmonton, AB, T6G 2G4
Canada
Phone: 1 7804920404
Email: cweiman@ualberta.ca

Abstract

Background: Military members (MMs) and public safety personnel (PSP) are vulnerable to occupational stress injuries because of their job demands. When MMs and PSP transition out of these professions, they may continue to experience mental health challenges. The development and implementation of resilience-building mobile health (mHealth) apps as an emergent mental health intervention platform has allowed for targeted, cost-effective, and easily accessible treatment when in-person therapy may be limited or unavailable. However, current mHealth app development is not regulated, and often lacks both clear evidence-based research and the input of health care professionals.

Objective: This study aims to evaluate the evidence-based quality, efficacy, and effectiveness of resilience-building mobile apps targeted toward the MMs, PSP, and veteran populations via a scoping literature review of the current evidence base regarding resilience apps for these populations and an evaluation of free resilience apps designed for use among these populations.

Methods: The studies were selected using a comprehensive search of MEDLINE, CINAHL Plus, PsycINFO, SocINDEX, Academic Search Complete, Embase, and Google and were guided by PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews). A narrative synthesis of the resulting papers was performed. The Alberta Rating Index for Apps was used to conduct a review of each of the identified apps. The inclusion criteria consisted of apps that were free to download in either the Google Play Store or the Apple App Store; updated within the last 3 years; available in English and in Canada; and intended for use by MMs, veterans, and PSP.

Results: In total, 22 apps met the inclusion criteria for evaluation. The resilience strategies offered by most apps included psychoeducation, mindfulness, cognitive behavioral therapy, and acceptance and commitment therapy. Overall, 50% (11/22) of apps had been tested in randomized controlled trials, 7 (32%) apps had been evaluated using other research methods, and 5 (23%) apps had not been studied. Using the Alberta Rating Index for Apps, the app scores ranged from 37 to 56 out of 72, with higher rated apps demonstrating increased usability and security features.

Conclusions: The mHealth apps reviewed are well-suited to providing resilience strategies for MMs, PSP, and veterans. They offer easy accessibility to evidence-based tools while working to encourage the use of emotional and professional support with safety in mind. Although not intended to function as a substitute for professional services, research has demonstrated that mHealth apps have the potential to foster a significant reduction in symptom severity for posttraumatic stress disorder, depression, anxiety,

XSL•FO
RenderX

and other mental health conditions. In clinical practice, apps can be used to supplement treatment and provide clients with population-specific confidential tools to increase engagement in the treatment process.

(JMIR Mhealth Uhealth 2022;10(1):e26453) doi:10.2196/26453

KEYWORDS

occupational stress injury; trauma; mHealth; resilience; mental health; military; veteran; public safety personnel; OSI; PTSD; mental health intervention; mobile phone

Introduction

Background

Globally, military members (MMs) and public safety personnel (PSP), for example, correctional workers, dispatchers, firefighters, paramedics, and police officers, experience increased exposure to trauma and stress in their daily activities, which can affect their mental health and well-being [1,2]. PSP and MMs are at an increased risk of developing occupational stress injuries (OSIs), including posttraumatic stress disorder (PTSD), major depressive disorder, generalized anxiety disorder, and increased anger, aggression, or hostility, which can lead to other challenges, such as substance abuse, relationship difficulties, and workplace absenteeism [1].

MMs in the Canadian Armed Forces (CAF) are at greater risk of mental health disorders and suicide risk compared with the Canadian civilian population [3]. Of the regular force CAF members, 32.2% self-reported a mental health problem related to emotions, stress, substances, or family in 2013-2014 [4]. In addition, within the 3-year period from 2013 to 2016, mental health conditions in the veteran populations showed an increase from 25.4% to 30.3%, with PTSD being the most commonly identified OSI [5]. A 2018 study indicated that across the Canadian PSP groups, 44% screened positive for at least one mental health disorder [1]. This study also found that 36.7% of surveyed Canadian police officers in particular screened positive for mental health conditions, primarily PTSD [1]. These populations face challenges related to attaining professional mental health, including displacement owing to relocation, working in remote geographic locations, and shift work.

Owing to the need for mental health support among MMs, veterans, and PSP, mobile health (mHealth) apps have emerged as a portable treatment modality option [6-8]. Interest and use of mHealth by clinicians has increased in recent years in health care practice [9]. The latest estimates suggest that there are between 165,000 and 325,000 health and wellness apps currently available for download [10,11]. When considering the MMs, veteran, and PSP populations, mHealth apps have gained popularity as a mental health treatment modality because of their low costs, easy access, and in-the-moment interventions [12].

Resilience

Evidence illustrates that resilience training and interventions, primarily those focused on coping skills and self-efficacy, can work to support a decrease in psychological distress and symptoms of PTSD [13-16]. Resilience is a broad and often complex concept, which scholars have uniquely interpreted depending on the context and can encompass both the individual and the group [17]. For this study, we have defined resilience as follows: “The dynamic process of overcoming adverse experiences through the use of internal and external resources in order to foster healthy psychological functioning” [13,17-19].

Resilience has been identified as an important factor that enables individuals adapt to and recover from emotionally, physically, and psychologically distressing situations and trauma [17,20]. There have been a multitude of resilience models and frameworks proposed in recent years specific to military, veteran, and PSP populations. One such model was developed using a large meta-analysis study by the Defence Human Capability and Science Technology Centre in 2014 [21]. This was further refined by Precious and Lindsay [22] with the Australian Armed Forces, resulting in a pillar of the mental resilience model (Figure 1 [21,22]). This model collaboratively draws on the best evidence related to mental resilience, highlighting both aspects outside one’s locus of control (ie, learned skills, previous experience, and personality) and the activities and skills within one’s locus of control (ie, mental control, emotional regulation, coping, self-efficacy, sense of purpose, positive affect, and social support) [22].

XSL•FO
RenderX

Figure 1. Pillars of mental resilience [21,22].

The activities and skills listed in this model have been attributed to the fostering of resilience, including those used to improve emotional regulation and coping, such as mindfulness, grounding, and self-talk; positive affect, such as purposeful leisure activities; and interpersonal relationships [23,24]. According to Lopez [23], "resilient individuals have a greater likelihood of engaging in healthy and productive activities and having a better quality of life." If an individual does not possess the self-regulatory abilities and tools necessary to deal with distressing situations, it can impact all their areas of life, including sleep, quality of life, work, motivation, interest, and engagement in daily activities [24]. Resilience is, therefore, vital for the MMs, veteran, and PSP populations, as it supports continued engagement in both purposeful activities and increases the ability to adapt to the challenges of daily living.

Digital Health

The term *digital health* refers to the use of electronic communication, services, and processes to deliver and facilitate health care services [25]. mHealth is a more recent subsegment of digital health that uses mobile technology to enable remote care and clinical health data collection. Digital health–based treatments have become a growing field of research and development for the MMs, PSP, and veteran populations. Popular modalities include virtual reality, web-based programs and games, and mHealth apps [26-28]. Current research indicates that these platforms are effective at reducing the symptoms of PTSD and other mental health disorders caused by exposure to trauma [28-32].

Specific to the MMs, PSP, and veteran populations, health care professionals (HCPs) are well-suited to use mHealth tools, such as apps, to supplement treatment and provide clients with immediate tools to help them overcome psychological impairment related to their traumas [9,31]. For example, if a military member or PSP encounters a stressful situation, the accessibility of apps can help them navigate their feelings in real time. In addition, mHealth apps may fill a gap for those who require mental health treatment but are faced with barriers, such as stigma to visible help-seeking, long waiting times, a high mobility of their jobs, and geographic restrictions [8,30,33,34], all of which have been noted as problematic for the MMs, PSP, and veteran populations. Many apps are also cost-effective and may be beneficial where therapy services are limited or unavailable [35,36].

Although mHealth tools have significant potential, several barriers limit their full uptake in the health care system. For example, some forms of technology, though widely used, may still not be available to everyone under all circumstances, for example, locations with unreliable or reduced cellular service, limited Wi-Fi access, and financial barriers may impede their use [37]. More importantly, there is a paucity of peer-reviewed research published regarding mHealth apps to determine their uptake, impact, or the best practices for their development and regulation [12,30,37]. When considering the use and development of mHealth apps, the limited number of research studies and agencies available to regulate this field impacts the ability to meet the current needs of this rapidly expanding industry [12,37]. As a result, a significant proportion of apps currently available for download have limited evidence of the effectiveness, efficacy, and safety of their use. This can make it difficult for HCPs to identify the best mHealth resources to recommend to MMs, PSP, and veterans in support of their care.

Objectives

The aim of this study is to evaluate the evidence-based quality, efficacy, and effectiveness of resilience-building mobile apps targeted toward the MMs, PSP, and veteran populations. This is addressed through two objectives: (1) completion of a scoping literature review of the current evidence base regarding mental health apps for these populations and (2) evaluation of common free mental health apps designed for use among these populations. We then compare and triangulate the data from these 2 approaches. The determination of these factors will aid in improving the evidence base for mHealth apps to highlight their potential use for HCPs who may be providing mental health services.

XSL•FO
RenderX

Methods

Objective 1

A Scoping Literature Review

A scoping literature review was completed to explore the available literature on mHealth apps and their cultivation of resilience in MMs, PSP, and veteran populations. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) was used to guide this scoping review and app search, both of which were conducted between December 18, 2020, and December 20, 2020 [38]. We selected the following electronic databases for the search: MEDLINE (Ovid interface), CINAHL Plus with Full Text (EBSCOhost interface), PsycINFO (Ovid interface), SocINDEX with Full Text (EBSCOhost interface), Academic Search Complete (EBSCOhost interface), and Embase (Ovid interface). We also used Google as a search tool to investigate gray literature in case it was not detected in the chosen database. Additional resources were manually selected to ensure a comprehensive retrieval of relevant studies that may have fallen outside of the predetermined search terms. The following steps were adhered to: (1) determination of the population, intervention, comparison, and outcomes research question, (2) determination of an eligibility criteria, (3) definition of search terms (Multimedia Appendix 1), (4) title and abstract screening, (5) full-text reading, (6) charting of the data, and (7) narrative synthesis. The final literature search took place on December 20, 2020.

Determination of Research Question

This literature review aimed to answer the following research question: What is the efficacy, effectiveness, and quality of mHealth apps on increasing resilience and self-regulatory strategies among MMs, PSP, and veterans?

Determination of Eligibility Criteria

The literature included in the search encompassed studies published from the year 2000 onward to account for the development of technology during this time. The articles had to address resilience or self-regulatory strategies. In addition, the articles were required to pertain to military, veteran, or PSP populations (Textbox 1).

Textbox 1. Eligibility criteria for scoping literature review.

Inclusion criteria

- The search was limited to studies published from the year 2000 onward to include more current technology.
- Included articles focused on participants aged ≥16 years.
- Articles addressing resiliency, hardiness, or coping.

Exclusion criteria

- Data not pertaining to military populations or public service personnel.
- Studies published in languages other than English.
- No outcome of interest.

Definition of Search Terms

Keywords for the search were determined using three main concepts: specific population, resilience, and games (refer to Multimedia Appendix 1 for a full description of the search terms).

Title and Abstract Screening

After the removal of duplicate articles from the search results, a minimum of 2 researchers screened each article based on their titles and abstracts to determine further eligibility for the literature review. Articles that did not meet the eligibility criteria were excluded. Conflicts were discussed and resolved via team consensus.

Full-Text Reads

The screened articles were then read in full by a minimum of 2 researchers. Conflicts were discussed and resolved via team meetings and final eliminations were made. The remaining articles were included in the scoping review.

Charting of the Data

The type of evidence, population, funding, interventions, outcomes, and recommendations were extracted from each remaining article and recorded on a spreadsheet.

Narrative Synthesis

A narrative synthesis was performed by 3 researchers to summarize the findings of the different studies and evaluation results. Narrative synthesis refers to an approach that relies primarily on the use of words and text to summarize and explain the findings of multiple studies associated with reviews [39]. Narrative synthesis can be particularly helpful when studies are heterogeneous and organizing the data in a more numerical or statistical format would be inappropriate. To conduct the narrative synthesis, 3 researchers first reviewed the included study results and deductively organized them using the pillars of mental resilience as a guide. Additional information related to study methods, key constructs, and study outcomes was then synthesized together to form a coherent understanding of each topic. Finally, the researchers provided a summary of the included articles and their relevance to app evaluations.

XSL•FO
RenderX

Objective 2: Evaluation of Available Mental Health Apps

The identified apps were chosen through the following steps: (1) identification of the apps addressed in the literature review, (2) establishment of eligibility criteria, and (3) search of the eligible apps by name in the Apple App Store or Google Play Store. Apps were then evaluated for overall quality using the Alberta Rating Index for Apps (ARIA) [40].

Eligibility Criteria

Apps included in the study were available for download in the Apple App Store or Google Play Store, could be set up in English, and were accessible within the geographic region of Canada. Apps chosen were intended for use primarily with MMs, PSP, and veterans; however, app use and availability could also extend to civilian populations. We included only free apps because evidence indicates that although 93% of smartphone users are likely to download an app, only 35.8% would be inclined to pay for an app [41]. Refer to Textbox 2 for the detailed eligibility criteria.

Textbox 2. Eligibility criteria for mobile health apps included in the study.

Inclusion criteria

- Apps that were available on the Apple App Store and Google Play Store.
- Apps that were free to download.
- Apps that were intended for use by military members or public safety personnel.

Exclusion criteria

- Apps that were not free to download.
- Apps that were not available in the English language.
- Apps that were not available in Canada.
- Apps that were not yet released for public use or access.

Outcome Measure: ARIA

The ARIA (Multimedia Appendix 2 [40]) was used as a measuring tool to rate each app included in the study [40]. There are two versions of the ARIA: one for care providers and one for end users. Although the ARIA has yet to be vigorously studied, its uptake by health care systems at the regional and national level has been swift, likely owing to its ease of use, applicability, and both clinical utility and accessibility to the client population. The ARIA is available in both English and French, allowing it to be used across Canada, particularly with all PSP and members of the CAF and Veterans Affairs Canada (VAC). The tool has 2 sections, A and B, with multiple feature items to be rated between 0 and 4: 0 being *strongly disagree* and 4 being *strongly agree*. Section A was completed before downloading the app. Items in section A included purpose, trustworthiness, privacy, and affordability [42]. Once completed, scores were added up for section A to obtain a total out of 24 [42]. Section B is scored once the user has spent a minimum of 10 minutes using the app [42]. The items in section B include security, trustworthiness, ease of use, functionality, target users, usefulness, and satisfaction [42]. When completed, the rater added up the section to obtain a score out of 48. Sections A and B were then added together to determine an overall score out of 72, with higher scores indicating better performance and user experience. The care-provider version also features a 0-4 rating scale on whether the app would be recommended to possible users. This measure does not impact the overall score of the app [42].

The care-provider version of the ARIA was chosen to review the selected apps because of its ability to be implemented by users, caregivers, and HCPs. Each identified app that met the eligibility criteria was evaluated by 2 reviewers using the ARIA. Conflicts were resolved through discussion with all researchers to determine the final ARIA score.

Results

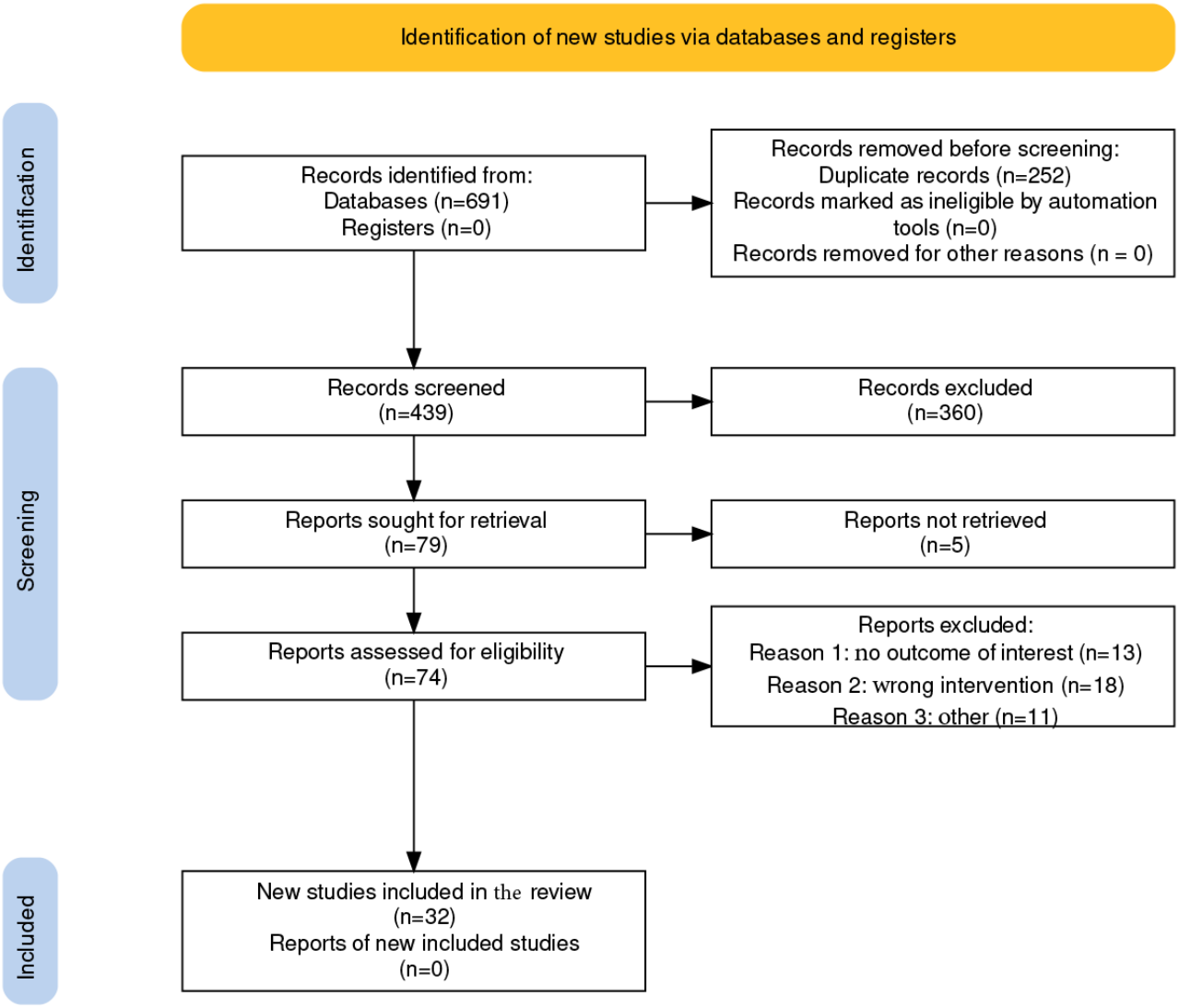
Objective 1: Scoping Literature Review

Study Selection

After searching the databases and identifying records through additional sources, a total of 691 articles were identified. Seven additional records were identified from the other sources. After the removal of 252 duplicates, 63.5% (439/691) of articles remained for title and abstract screening. A total of 52.1% (360/691) of articles were determined to be irrelevant. Full-text reads were completed on the remaining 10.7% (74/691) of articles and 6.1% (42/691) of additional articles were excluded because of differences in outcomes of interest, irrelevancy to mHealth app use, limited qualitative or quantitative data, and repetitive publications. A total of 9 studies were excluded because they did not have a relevant outcome of interest. The reasons for exclusion were that the studies involved virtual reality without an app and were specifically for outcomes related to PTSD. From these 32 articles, 22 apps were identified as meeting the inclusion criteria. Figure 2 shows a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) summary chart of the study selection process. The results of the scoping review, including narrative synthesis, are described in subsequent sections.

XSL•FO
RenderX

Figure 2. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) chart for the systematic review study identification, selection, exclusion, and inclusion.



Study Findings

Evidence-Based Merit

Of the 22 apps identified in the scoping literature review, 11 (50%) apps had been tested in randomized controlled trials (RCTs), 7 (32%) apps using other methods, and 5 (23%) apps had not been research trialed at the time of this study. Alternative methods used to evaluate apps included a nonrandomized quasi-experimental pre–post follow-up design, pre- and posttest questionnaires, qualitative focus groups, and the Mobile App Rating Scale. Apps that were not research trialed were identified through scoping reviews. Two apps had been tested using multiple methods.

The literature indicated that out of the 22 apps, 15 (68%) used evidence-based strategies or incorporated evidence-based components within them. Among these 15 apps, 11 (73%) apps were developed using evidence-based practices as their foundation (*Virtual Hope Box*, *eQuoo*, *Mindfulness Coach*, *Resilience@Work/Mindarma*, *PE Coach*, *R2MR*, *High Res*, *PTSD Coach*, *CBT-I Coach*, *PHIT for Duty*, and *Stay Quit Coach*). In the RCT by Roy et al [28], 7 apps were included in determining the effectiveness of symptom reduction (*Positive Activity Jackpot*, *Tactical Breather*, *Daily Yoga*, *Simply Yoga*, *Life Armor*, *PE Coach*, and *Virtual Hope Box*). However, it should be noted that apps were not separated but rather assessed as a group, rendering it difficult to identify the efficacy and effectiveness of each app individually. *Virtual Hope Box*, *eQuoo*, and *Resilience@Work/Mindarma* were evaluated separately in their respective RCT studies.

Mental Control, Emotional Regulation, Coping, and Self-efficacy

The resilience strategies offered by most of the apps fit within the pillars of mental resilience, including mindfulness training, psychoeducation, cognitive behavioral therapy, and acceptance and commitment therapy. Other strategies included biofeedback, sleep strategies, social engagement, mood tracking, time scheduling, and muscle relaxation techniques, such as yoga. Many of the apps included more than one strategy. Of the total 22 apps, 8 (36%) apps used mindfulness strategies, including *Resilience@Work/Mindarma*, *Mindfulness Coach*, *Daily Yoga*, *Simply Yoga*, *Virtual Hope Box*, *Tactical Breather*, *Breathe2Relax*, and *PHIT for Duty* and 4 (18%) apps

XSL•FO
RenderX

encompassed psychoeducation: *eQuoo*, *Resilience@Work/Mindarma*, *Life Armor*, and *PE Coach*. Apps that applied cognitive behavioral therapy techniques included *R2MR*, *High Res*, *PTSD Coach*, *Family Coach*, *CBT-I Coach*, *Stay Quit Coach*, and *eQuoo*. *Resilience@Work/Mindarma* was the only app in this study that drew from acceptance and commitment therapy principles.

Effect of Apps on Resilience

The evidence-based literature demonstrated that many of the apps increased resilience strategies for users as well as improved the overall aspects of mental health [28,30,43,44]. The *eQuoo* app was demonstrated to significantly improve the traits of resilience, personal growth, and positive relationships [40]. Similarly, *Resilience@Work/Mindarma* showed improved adaptive resilience and psychological flexibility [15]. Joyce et al [15] also found that this app significantly increased optimism and mindfulness practice among study participants. This study also found that the app increased use of emotional support from others and help-seeking behavior, which addresses the social support pillar of the pillars of mental resilience. In addition, *R2MR*, which stands for road to mental readiness, was found to be effective in increasing resilience and help-seeking behaviors in the participants and in reducing mental health stigma for individuals and entire workplaces [45]. Although much of the literature noted the integration of social support as a positive influence on app use, resilience and coping, and health promotion, there was an observed lack of social connectedness components within the apps reviewed [14,23,36,46].

Health Care and Social Support

Some apps, such as *PTSD Coach*, were demonstrated to be more effective in managing and reducing PTSD symptoms when used with the support of an HCP as opposed to independent use [43,47]. Sessions that provide instruction around optimal app use patterns as well as the app's purpose, can increase the user's knowledge and therefore adherence to treatment programs that uses mHealth components [43,48,49]. Additional mentoring or coaching may also contribute to a greater elaboration on techniques introduced within apps and may enable the transfer of skills from the app experience [14]. Evidence further indicates that using mHealth apps with support as supplementary resources, rather than primary treatment, may enhance therapeutic outcomes and allow users more autonomy in their ability to track symptoms while sharing results with their providers [14,30,43]. When users complete treatment sessions, they have the ability to retain these tools for future use or reference to their care [30,50].

End User Preferences, Incentives, and Real-world Apps

Within the studies, there were important themes that arose through narrative synthesis around user preferences. One of those themes included apps that had a sense of progression, rhythm and routine, and elements of personal causation [43]. Having set challenges and a clear visualization of the progress helped increase app use, adherence to the intervention, and goal attainment, especially when users were able to establish their targeted goals beforehand [26,51]. Earning rewards increased the attractiveness of the app, as did receiving guidance and instant feedback on target behaviors [26]. Apps designed around games and narratives were often preferred as they encompassed many of these traits and were user-friendly and enjoyable [51]. However, many users also noted a preference for more practical application opportunities of target skills, so their learned behaviors could be transferred to real-world concerns [52].

Objective 2: Evaluation of Mental Health Apps

The ARIA scores ranged from 37 to 56 out of 72 (Multimedia Appendix 3). The highest overall scoring apps were *R2MR*, *PTSD Coach*, and *AIMS for Anger Management*, all of which had a total score of 56 (Figure 3). The lowest scoring apps were *Resilience@Work/Mindarma* (37 out of 72), *Breathe2Relax* (38 out of 72), and *High Res* (39 out of 72).

Figure 3. Highest scoring apps on Alberta Rating Index for Apps.

Certain items of the ARIA were identified as being particularly relevant to the MMs and PSP populations. These included security and confidentiality, trustworthiness, and usefulness and satisfaction. As security and confidentiality are a high priority for the MMs and PSP populations, the security item of the ARIA is imperative to acknowledge [30]. The following apps rated the highest on the security items with a 4 out of 4 rating on both security and consent: *Virtual Hope Box*, *Positive Activity Jackpot*, *Daily Yoga*, and *Life Armor*. The apps that were rated 1 or less were *eQuoo*, *Resilience@Work/Mindarma*, *High Res*, and *PHIT for Duty*. Many apps were missing features such as a password or biometric identifier and instead relied on the user's phone security setup to provide these privacy measures.

XSL•FO
RenderX

Apps were rated higher on trustworthiness if developed by reliable sources with proof of evidence. In total, 15 apps were developed by government agencies in Canada and the United States, including the Department of Veteran Affairs, the Department of Defence, the National Centre for Telehealth and Technology, and VAC. For trustworthiness, *PTSD Family Coach*, *R2MR*, *Stay Quit Coach*, *CBT-i Coach*, and *eQuoo*, all scored 4 out of 4. Many apps did not state risk warnings associated with app use directly on their download page; however, information could occasionally be found within the app itself. For these cases, the apps were rated 1 out of 4 in trustworthiness. In the usefulness and satisfaction items, apps that rated the highest included *Mindfulness Coach* and *PTSD Coach*, both with combined scores of 11 out of 16 for the sections.

Discussion

Summary of Evidence

The aim of this study is to evaluate the evidence-based quality, efficacy, and effectiveness of resilience-building mobile apps targeted toward the MMs, PSP, and veteran populations. This involved the completion of a scoping literature review of the current evidence base regarding mental health apps for these populations and the evaluation of common free mental health apps designed for use among these populations. This study aims to provide some insight into the following research question: What is the efficacy, effectiveness, and quality of mHealth apps on increasing resilience and self-regulatory strategies among MMs, PSP, and veterans?

Overall, the results of this study indicated that most of the mHealth apps reviewed were well-suited to provide resilience strategies and skills for MMs, PSP, and veterans. These apps provided skills, strategies, and services, which could be categorized into the pillars of mental resilience and other commonly accepted definitions regarding psychoeducational interventions that can foster resilience. Common resilience strategies were well represented in many of the apps, often including mindfulness, psychoeducation, and positive coping or thinking skills. Our results indicated that no app fully addressed the 6 pillars of resilience identified by the armed forces of the United Kingdom and Australia. In total, 5 apps addressed 5 of the 6 pillars, whereas 12 apps addressed 4 or more of the 6 pillars, and 20 apps addressed 2 or more of the 6 pillars. The apps that rated the highest on the ARIA were *R2MR*, *Virtual Hope Box*, *eQuoo*, *Mindfulness Coach*, *and PTSD Family Coach*. The pillar most likely to be missing was *socialsupport*, with the apps largely ignoring this concept.

In the evaluation of the 22 apps, *R2MR*, *PTSD Coach*, and *AIMS for Anger Management* had the highest overall scores. Points were generally lost because of a missing statement of the risk of use of the app or a lack of security measures to protect app access from the individual's phone. Of the 22 apps assessed, 15 (68%) apps were developed by credible, military-focused government agencies (eg, VAC and the US Department of Veterans Affairs), which may help ensure that the content delivered was well-adapted for these populations. All but 2 of these apps were developed outside of Canada, which may impact the accessibility to local resources and services owing to the geographically based content (eg, helplines). Future comparisons of the ARIA with other app evaluation tools, such as the Mobile App Rating Scale, may allow for a more in-depth understanding of mHealth apps; however, for the MMs and PSP populations, the ARIA's additional security and privacy questions provide a clearer understanding of population-specific concerns [7,44].

When considering evidence around the apps, it was noted that out of the 22 apps, only 11 (50%) apps had undergone evidence-based evaluation through an RCT. Although from the total 22 apps, only 11 (50%) apps were determined to be evidence-based, 15 (68%) apps had used evidence-based strategies or components within them. Some of the apps selected for evaluation in this study have not been evaluated in the evidence-based literature. This is partially a result of the large creation and turnover of mHealth apps available for download as well as the currently limited regulations guiding their development [7,12,30]. With this being an understudied area of research, a lack of evidence influenced other potentially important client considerations. For example, parameters around effective dosages of apps, such as how long and how often a user was required to use the tool to see lasting effects, were not addressed. Instead, many of the apps identified in the studies relied on user feedback to conclude whether the application was clinically effective [50]. Another usability component considered through the ARIA was the presentation of information. Many of the apps relied on large blocks of text to present educational information and, as noted by O'Toole and Brown [30], this format can be overwhelming for users and cause disadvantages for those with alternative learning styles. This review illustrates the limitations of both the evidence and the potential quality of the apps being proposed to support resilience in MMs, PSP, and veterans.

Although these apps incorporate strategies and skills that may assist in facilitating resilience, it must be acknowledged that there are other factors that impact their ability to increase foster resilience, such as individual motivation, education on use of the app, access to social support, and the use of apps together with an HCP or independently. In addition, care should be taken regarding the specific designation of these apps as *resilience apps* (particularly in light of the MMs, PSP, and veteran populations). As a universal operational definition of resilience is lacking, a clear understanding of what elements constitute or are most important to resilience is also lacking. To illustrate, the Canadian studies of resilience in MMs have independently and not always cohesively explored the constructs of personality, positive affect, mastery, and social support [53]; neuroticism, military hardiness, and problem-solving coping [54]; and conscientiousness, emotional stability, and positive social interactions [55]. This inability to effectively define what is meant by resilience within a specific organization, such as the CAF, is problematic for both the research and intervention development. This problem is compounded when exploring how the construct is defined by the militaries of other countries (ie, Australia, United Kingdom, and United States vs Canada) and how it may enhance mental health [56]. Without research to definitively measure resilience before and after app use and without an effective definition of resilience—which impairs the

XSL•FO
RenderX

researcher's ability to quantify the concept—it is not possible to decisively conclude whether these apps increase resilience.

The Role of HCPs in mHealth and Resilience

Mental health can impact the daily functioning of an individual, and the concept of resilience is closely tied to mental health and well-being [23]. MMs and PSP are more likely to be exposed to traumatic experiences and are well-suited to resilience interventions [4,5]. HCPs can support individuals in navigating the environment for external resources and addressing the barriers in multiple domains that they may be experiencing.

Smartphones and technology are part of daily habits in the modern era, and HCPs can identify how to incorporate mHealth apps into health care settings. As MMs, PSP, and veterans face many unique challenges in terms of sudden environmental, lifestyle, and role changes, mHealth tools can present a more feasible option for access [24]. In a clinical context, HCPs can recommend apps to provide ongoing support outside of therapy services, create a tool for sharing health information, increase engagement in the treatment process, and sustain benefits gained once the provided services end [28]. For the mHealth apps themselves, the inclusion of both clinicians and users in their development can both ensure strategies meet the needs of their clients, thus encouraging wider acceptability and utility and helping clinicians better identify evidence-based features [30,57,58].

HCPs have a responsibility to advocate for best practices; this can be challenging with mHealth because of the high rate of app development and the inability for evidence-based practice to keep up [12,30,37]. Using tools such as the ARIA can help clinicians determine an app's usability and evidence base; however, HCPs must also ensure that the app is a good fit for clients in terms of their interests, values, abilities, and routines. HCPs can then use this information to collaborate with the client, customize apps that meet the client's needs and interests, and promote engagement with the apps. Creating client autonomy through app literacy will allow users to take more control in choosing treatment methods and encourage greater client-centered practice [59]. It is important to note, however, that many of these apps perform best when combined with the services of HCPs. This may include providing learning sessions before use to increase efficacy and optimal use patterns, weekly phone check-ins, or using the apps to enhance existing therapeutic interactions between clinicians and users [43,48-50]. In much of the literature, gamification and the integration of social components strengthened mHealth app use and engagement, resulting in more positive outcomes and an increased sense of peer support for the MMs and PSP populations [14,36,46]; however, this was lacking in practice.

Strengths and Limitations

This study has several strengths. Both the scoping literature review and app evaluation were conducted following a planned a priori procedure, with attention to ensuring quality control and minimizing bias. The detailed search strategy in the literature review was extensive, including 6 databases. The inclusion and exclusion criteria were determined before the study onset and adhered to throughout for both the literature review and app evaluation. We also used appropriate calibration and at least two independent reviewers for all stages of the process.

There are certain limitations to this study, which should also be acknowledged. In the literature review, only studies written in English were included. The app selection criteria were limited to apps available for download in Canada, which excluded potentially beneficial apps available in other geographic locations. In addition, as the researchers only had access to iPhones, the apps were not tested on Android devices, which could have an impact on the usability criteria.

Although the authors identified the ARIA as an appropriate evaluation tool to use, it should be acknowledged that this is still relatively new and has not yet been extensively researched, used, or validated at this point. As such, there are currently no similar studies conducted with the ARIA, with which the present results could be compared. The results listed within this study only reflect app use from the perspective of clinicians, which could create bias. It will be important in future studies to invite users from the MMs, PSP, and veteran populations to complete the *user version* of the ARIA.

Future Research and Directions

Regarding the future of mHealth and resilience, there is much work to do in the areas of research, development, and policy. First, despite its use in health care contexts, future research is required to determine how the ARIA scores correlate to app adherence, acceptance, and adoption by users as well as to health outcomes. Further comparison of the ARIA with other app evaluation tools would be valuable in understanding the utility, criterion validity, and other concepts related to its ability to rate apps from the perspective of HCPs and clients as end users.

As previously mentioned, there is a paucity of evidence-based studies on the existing apps geared toward both resilience and the specific patient population of MMs, PSP, and veterans. Although this lack of research does not necessarily indicate that apps are of poor quality, it highlights the need for further research on health app development to ensure safety, effectiveness, and efficacy. It has been identified that traditional study design and research methods may be inappropriate for the study of mHealth as technology evolves much faster than traditional evidence-based research [8,60]. The lack of empirical research to demonstrate the effectiveness of apps on resilience may be related to the short time frame during which mHealth apps have emerged and the speed at which their availability changes [8]. Innovative and novel research methods that can address the demands of mHealth and the rapidly changing world of app development and use are needed to assist with the quality control of these tools used by both HCPs and patient populations.

Another area of study related to mHealth that is important and specific to military and PSP users is privacy, security, and confidentiality. Although there is a higher expectation of privacy for apps that involve health care information, military and PSP organizations may be subject to other restrictions on internet access that may impede the use of the app or demand higher security. Data sharing and privacy are considerations that require

XSL•FO
RenderX

attention from researchers, HCPs, and the general public when deciding on which app to use or if app use is appropriate at all [8]. Future systems of app evaluation and research would benefit from adding a component that considers data sharing, storage, and privacy in highly sensitive and secure patient populations (ie, military and PSP).

Future initiatives to assist HCPs and their clients in navigating the world of mHealth would be an asset to balance client autonomy through app literacy and would assure that apps have some level of evidence-based merit. As mHealth use is on the rise among many HCP and client populations, training to use apps to support service delivery is of utmost importance for health care organizations. The establishment of clear practice guidelines is important for both HCPs and clients, so that expectations about the usefulness and effectiveness of the app are appropriately managed. Currently, clients may have overly enthusiastic ideas about the effectiveness of these apps to develop their resilience and support their mental health, even when seeking professional mental health treatment might be necessary. Similarly, issues of risk and safety in mental health apps (including resilience) also need to be addressed.

The question of what constitutes a resilience app versus, for example, a wellness app, is also a murky territory that requires further navigation. Until a universally operationalized definition has been established for resilience, it is likely that confusion will remain regarding the codification of specific resilience skills and strategies. For example, family- and community-level factors were seldom addressed in our identified apps, despite being indicated as an important component of resilience [44,46]. This lack of inclusion is more surprising given the strong evidence that social support is a strong contributor to psychological health [23,24], overall well-being, and quality of life [24]. Similarly, empowering others to use their skills for stress reduction, coping, and building self-efficacy has been demonstrated to foster a significant reduction in the severity of the symptoms of PTSD, depression, and anxiety [12,13,28,58], which also alludes to the idea that the identified apps could be classified in terms of specific mental disorders. Until researchers, app developers, the various military and PSP organizations, and HCPs can agree on the definitions of these concepts, determining which intervention targets and affects resilience will remain elusive.

Finally, the evaluation of resilience apps will remain challenging for all stakeholders and can affect the quality of the product unless all stakeholders were included in the consultation process. With the exception of the apps designed by the United States Department of Defense of Veterans Affairs, it was difficult to determine if the end user's perspective was incorporated into the development of the app. A collaborative approach to development, using both expert and user input, has been noted in recent studies as an effective approach to increasing the success of apps [57,58,60]. Ideally, in the future, mHealth development should engage the end users' input to assist with contextualization, which may increase the app acceptance, usability, and feasibility in a multitude of health care and possibly military or PSP settings.

Conclusions

Resilience is often targeted by HCPs through interventions that strengthen social support systems, foster greater self-concept, encourage optimism, promote the ability to reflect, and build emotional strength. Although not intended to function as a substitute for professional services and interventions, mHealth apps have the potential to foster resilience and support a significant reduction in symptom severity for OSIs, including PTSD, depression, and anxiety, in populations affected by OSIs, such as MMs, veterans, and PSP. Apps provide easy accessibility to evidence-based tools and encourage users to initiate help-seeking behaviors when stigma or uncertainty may impede the use of direct care. In clinical practice, HCPs can assist clients in identifying apps that support their habits and values and bolster participation and engagement in activities of daily living. As accessible, novel, and evidence-based interventions and resources for fostering resilience and addressing mental health become available, MMs, veterans, and PSP may be able to facilitate their healing, recovery, and growth, which would have a positive effect on their families, communities, organizations, and the public they serve.

Acknowledgments

The authors wish to thank Dr Peyman Azad Khaneghah for offering the use of the Alberta Rating Index for Apps and providing support for its implementation. This study would not have been possible without the support of the Heroes in Mind, Advocacy, and Consortium research laboratory.

The Recipient is providing this report of the project results on an "as is, where is" basis and makes no representations or warranties, either express or implied, as to any matter including, without limitation, whether the project results or any part or aspect of the same will be capable of statutory protection, the existence or non-existence of competing technology, the condition, quality or freedom from error of the project results or any part thereof, any merchantability, or its fitness for any particular purpose and all warranties and conditions expressed or implied, statutory or otherwise are hereby disclaimed. Neither the Recipient nor its officers, directors, employees, students or agents will be liable for any direct, consequential or other damage suffered by anyone resulting from the development or use of the project results or any inventions, technology or product produced in the course of or using the project results. The user of this report and/or any project results contained in the report uses the report and/or project results at the user's own risk.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature search strategy for scoping review.
[DOCX File , 14 KB - mhealth_v10i1e26453_app1.docx]

Multimedia Appendix 2
Alberta Rating Index for Apps care-provider and user versions.
[PDF File (Adobe PDF File), 733 KB - mhealth_v10i1e26453_app2.pdf]

Multimedia Appendix 3
Alberta Rating Index for Apps score table.
[PDF File (Adobe PDF File), 42 KB - mhealth_v10i1e26453_app3.pdf]

References

1. Carleton RN, Afifi TO, Turner S, Taillieu T, Duranceau S, LeBouthillier DM, et al. Mental disorder symptoms among public safety personnel in Canada. Can J Psychiatry 2018 Jan;63(1):54-64 [FREE Full text] [doi: 10.1177/0706743717723825] [Medline: 28845686]
2. Ricciardelli R, Czarnuch S, Carleton RN, Gacek J, Shewmake J. Canadian public safety personnel and occupational stressors: how PSP interpret stressors on duty. Int J Environ Res Public Health 2020 Jul 01;17(13):4736 [FREE Full text] [doi: 10.3390/ijerph17134736] [Medline: 32630259]
3. Rusu C, Zamorski MA, Boulos D, Garber BG. Prevalence comparison of past-year mental disorders and suicidal behaviours in the Canadian armed forces and the Canadian general population. Can J Psychiatry 2016 Apr;61(1 Suppl):46S-55S [FREE Full text] [doi: 10.1177/0706743716628856] [Medline: 27270741]
4. Theriault F, Gabler K, Naicker K. Health and lifestyle information survey of Canadian forces personnel 2013/2014. Regular Force Report. 2016. URL: https://www.canada.ca/content/dam/dnd-mdn/documents/health/health-and-lifestyle-survey-2013-2014.pdf [accessed 2021-12-23]
5. Thompson J, Van Til L, Poirier A, Sweet J, McKinnon K, Pedler D. Health and well-being of Canadian armed forces veterans: findings from the 2013 life after service survey. Research Directorate, Veterans Affairs Canada. 2014. URL: https://www.veterans.gc.ca/pdf/about-us/research-directorate/2013-survey-caf-health.pdf [accessed 2021-12-23]
6. Beshai S, Carleton N. Peer support and crisis-focused psychological intervention programs in Canadian first responders: blue paper. University of Regina Collaborative Centre for Justice and Safety. 2016. URL: https://www.cipsrt-icrtsp.ca/en/publication/peer-support-and-crisis-focused-psychological-intervention-programs-in-canadian-first-responders-blue-paper [accessed 2021-12-23]
7. Tam-Seto L, Wood VM, Linden B, Stuart H. Perceptions of an AI-supported mobile app for military health in the Canadian armed forces. Military Behavioral Health 2020 Nov 13;9(3):247-254. [doi: 10.1080/21635781.2020.1838364]
8. Jones C, O'Toole K, Jones K, Brémault-Phillips S. Quality of psychoeducational apps for military members with mild traumatic brain injury: an evaluation utilizing the mobile application rating scale. JMIR Mhealth Uhealth 2020 Aug 18;8(8):e19807 [FREE Full text] [doi: 10.2196/19807] [Medline: 32808937]
9. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
10. Carlo AD, Hosseini Ghomi R, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019;2:54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
11. Lewis TL, Boissaud-Cooke MA, Aungst TD, Eysenbach G. Consensus on use of the term "App" versus "Application" for reporting of mHealth research. J Med Internet Res 2014 Jul 17;16(7):e174; discussion e174 [FREE Full text] [doi: 10.2196/jmir.3460] [Medline: 25033233]
12. Kuhn E, Owen JE. Advances in PTSD treatment delivery: the role of digital technology in PTSD treatment. Curr Treat Options Psych 2020 Mar 27;7(2):88-102. [doi: 10.1007/s40501-020-00207-x]
13. Combat-related trauma, posttraumatic stress disorder, and moral injury in combat medics: a comprehensive survey of the literature and critical analysis. ProQuest. URL: https://www.proquest.com/openview/ae7c04d8fdea0b1e660499b5ff4f8589/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2021-12-23]
14. de Visser EJ, Dorfman A, Chartrand D, Lamon J, Freedy E, Weltman G. Building resilience with the Stress Resilience Training System: design validation and applications. WOR 2016 Jul 05;54(2):351-366. [doi: 10.3233/wor-162295]
15. Joyce S, Shand F, Lal TJ, Mott B, Bryant RA, Harvey SB. Resilience@Work mindfulness program: results from a cluster randomized controlled trial with first responders. J Med Internet Res 2019 Feb 19;21(2):e12894 [FREE Full text] [doi: 10.2196/12894] [Medline: 30777846]

XSL•FO
RenderX

16. Pallavicini F, Argenton L, Toniazzi N, Aceti L, Mantovani F. Virtual reality applications for stress management training in the military. Aerospace Med Human Performance 2016 Dec 01;87(12):1021-1030. [doi: 10.3357/amhp.4596.2016]
17. Southwick SM, Bonanno GA, Masten AS, Panter-Brick C, Yehuda R. Resilience definitions, theory, and challenges: interdisciplinary perspectives. Eur J Psychotraumatol 2014 Oct;5 [FREE Full text] [doi: 10.3402/ejpt.v5.25338] [Medline: 25317257]
18. Masten AS. Resilience from a developmental systems perspective. World Psychiatry 2019 Feb;18(1):101-102 [FREE Full text] [doi: 10.1002/wps.20591] [Medline: 30600628]
19. Readiness, resilience, growth. Government of Canada. URL: https://army.gc.ca/en/3-canadian-division/readiness-resilience-growth.page [accessed 2021-12-23]
20. Green KT, Hayward LC, Williams AM, Dennis PA, Bryan BC, Taber KH, Mid-Atlantic Mental Illness Research, EducationClinical Center Workgroup, et al. Examining the factor structure of the Connor-Davidson Resilience Scale (CD-RISC) in a post-9/11 U.S. military veteran sample. Assessment 2014 Aug 27;21(4):443-451 [FREE Full text] [doi: 10.1177/1073191114524014] [Medline: 24586090]
21. Examination of the evidence for preparing, sustaining and enhancing psychological and physical well-being - psychological resilience in the UK Armed Forces. Institute for Employment Studies. 2014. URL: https://www.employment-studies.co.uk/project/examination-evidence-preparing-sustaining-and-enhancing-psychological-and-physical-well [accessed 2021-12-23]
22. Precious D, Lindsay A. Mental resilience training. J R Army Med Corps 2019 Apr 22;165(2):106-108. [doi: 10.1136/jramc-2018-001047] [Medline: 30580281]
23. Lopez A. Posttraumatic stress disorder and occupational performance: building resilience and fostering occupational adaptation. Work 2011;38(1):33-38. [doi: 10.3233/WOR-2011-1102] [Medline: 21248418]
24. Edgelow MM, MacPherson MM, Arnaly F, Tam-Seto L, Cramm HA. Occupational therapy and posttraumatic stress disorder: a scoping review. Can J Occup Ther 2019 Apr;86(2):148-157. [doi: 10.1177/0008417419831438] [Medline: 31014080]
25. Benefits of digital health: what is digital health? Canada Health Infoway. URL: https://www.infoway-inforoute.ca/en/what-we-do/benefits-of-digital-health/what-is-digital-health [accessed 2021-06-23]
26. Pusey M, Wong KW, Rappa NA. Resilience interventions using interactive technology: a scoping review. Interactive Learn Environ 2020 Jun 11:1-16. [doi: 10.1080/10494820.2020.1772837]
27. Rizzo A, Buckwalter JG, Forbell E, Reist C, Difede J, Rothbaum BO, et al. Virtual reality applications to address the wounds of war. Psychiatric Annals 2013 Mar;43(3):123-138. [doi: 10.3928/00485713-20130306-08]
28. Roy MJ, Costanzo ME, Highland KB, Olsen C, Clayborne D, Law W. An app a day keeps the doctor away: guided education and training via smartphones in subthreshold post traumatic stress disorder. Cyberpsychol Behav Soc Netw 2017 Aug;20(8):470-478. [doi: 10.1089/cyber.2017.0221] [Medline: 28737954]
29. Bisson JI, van Deursen R, Hannigan B, Kitchiner N, Barawi K, Jones K, et al. Randomized controlled trial of multi-modular motion-assisted memory desensitization and reconsolidation (3MDR) for male military veterans with treatment-resistant post-traumatic stress disorder. Acta Psychiatr Scand 2020 Aug;142(2):141-151. [doi: 10.1111/acps.13200] [Medline: 32495381]
30. O'Toole K, Brown CA. Evaluating the quality of resilience apps for military members and public safety personnel. J Military Veteran Family Health 2021 Feb 01;7(1):87-101. [doi: 10.3138/jmvfh-2020-0002]
31. van Gelderen MJ, Nijdam M, Haagen J, Vermetten E. Interactive motion-assisted exposure therapy for veterans with treatment-resistant posttraumatic stress disorder: a randomized controlled trial. Psychother Psychosom 2020;89(4):215-227 [FREE Full text] [doi: 10.1159/000505977] [Medline: 32203971]
32. mHealth App Developer Economics 2017: the current status and trends of the mHealth app market. Research2Guidance. URL: https://research2guidance.com/ [accessed 2021-12-23]
33. Bush NE, Smolenski DJ, Denneson LM, Williams HB, Thomas EK, Dobscha SK. A virtual hope box: randomized controlled trial of a smartphone app for emotional regulation and coping with distress. Psychiatr Serv 2017 Apr 01;68(4):330-336. [doi: 10.1176/appi.ps.201600283] [Medline: 27842473]
34. Vogt D. Mental health-related beliefs as a barrier to service use for military personnel and veterans: a review. Psychiatr Serv 2011 Feb;62(2):135-142. [doi: 10.1176/ps.62.2.pss6202_0135] [Medline: 21285091]
35. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
36. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
37. Imlawi J. Health website success: user engagement in health-related websites. Int J Interact Mob Technol 2017 Nov 27;11(6):49. [doi: 10.3991/ijim.v11i6.6959]
38. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

XSL•FO
RenderX

39. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product From the ESRC Methods Programme. Lancaster: Lancaster University; 2006.
40. Azad-Khaneghah P, Roduta Roberts M, Stroulia E, Ferguson-Pell M, Liu L. Alberta Rating Index for Apps (ARIA): helping older adults find acceptable mobile health apps. Gerontechnol 2020 Oct 03;19:1-1. [doi: 10.4017/gt.2020.19.s.69884]
41. Kertz S, MacLaren Kelly J, Stevens K, Schrock M, Danitz S. A review of free iPhone applications designed to target anxiety and worry. J Technol Behav Sci 2017 Jan 11;2(2):61-70 [FREE Full text] [doi: 10.1007/s41347-016-0006-y]
42. Azad Khaneghah P. Alberta Rating Index for Apps (ARIA): an index to rate the quality of mobile health applications. University of Alberta Education & Research Archive. 2020. URL: https://era.library.ualberta.ca/items/3c62ca2d-d7a1-4808-b1c0-4a60b05eecaa [accessed 2021-12-23]
43. Reyes AT, Serafica R, Sojobi A. College student veterans' experience with a mindfulness- and acceptance-based mobile app intervention for PTSD: a qualitative study. Arch Psychiatr Nurs 2020 Dec;34(6):497-506. [doi: 10.1016/j.apnu.2020.09.005] [Medline: 33280672]
44. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
45. Szeto A, Dobson KS, Knaak S. The road to mental readiness for first responders: a meta-analysis of program outcomes. Can J Psychiatry 2019 Jun;64(1_suppl):18S-29S [FREE Full text] [doi: 10.1177/0706743719842562] [Medline: 31010293]
46. Litvin S, Saunders R, Maier MA, Lüttke S. Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: a randomized controlled trial. PLoS One 2020;15(9):e0237220 [FREE Full text] [doi: 10.1371/journal.pone.0237220] [Medline: 32877425]
47. Possemato K, Kuhn E, Johnson E, Hoffman JE, Owen JE, Kanuri N, et al. Using PTSD Coach in primary care with and without clinician support: a pilot randomized controlled trial. Gen Hosp Psychiatry 2016;38:94-98. [doi: 10.1016/j.genhosppsych.2015.09.005] [Medline: 26589765]
48. Roy M, Highland KB, Costanzo MA. GETSmart: Guided Education and Training via Smart phones to promote resilience. Stud Health Technol Inform 2015;219:123-128. [Medline: 26799892]
49. Weltman G, Lamon J, Freedy E, Chartrand D. Police department personnel stress resilience training: an institutional case study. Glob Adv Health Med 2014 Mar;3(2):72-79 [FREE Full text] [doi: 10.7453/gahmj.2014.015] [Medline: 24808985]
50. Tam-Seto L, Wood VM, Linden B, Stuart H. A scoping review of mental health mobile apps for use by the military community. Mhealth 2018;4:57 [FREE Full text] [doi: 10.21037/mhealth.2018.12.01] [Medline: 30701175]
51. Puskar K, Sun R, Gleeson A, Lampi T, Nichols D, Khan N. MilTeenChat app to promote coping resilience in military youth. J Mil Veterans Health Jan 2018;26(1):44-46. [doi: 10.3316/informit.553541068174068]
52. Thomas KH, Taylor SP, Hamner K, Glazer J, Kaufman E. Multi-site programming offered to promote resilience in military Veterans. Calif J Health Promot 2015 Sep 01;13(2):15-24. [doi: 10.32398/cjhp.v13i2.1820]
53. Lee JE, Sudom KA, McCreary DR. Higher-order model of resilience in the Canadian forces. Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement 2011 Jul;43(3):222-234. [doi: 10.1037/a0024473]
54. Skomorovsky A, Stevens S. Testing a resilience model among Canadian forces recruits. Military Medicine 2013 Aug;178(8):829-837. [doi: 10.7205/milmed-d-12-00389]
55. Lee JE, Sudom KA, Zamorski MA. Longitudinal analysis of psychological resilience and mental health in Canadian military personnel returning from overseas deployment. J Occup Health Psychol 2013 Jul;18(3):327-337. [doi: 10.1037/a0033059] [Medline: 23834447]
56. van der Meulen E, van der Velden PG, van Aert RC, van Veldhoven MJ. Longitudinal associations of psychological resilience with mental health and functioning among military personnel: a meta-analysis of prospective studies. Soc Sci Med 2020 Jun;255:112814. [doi: 10.1016/j.socscimed.2020.112814] [Medline: 32388075]
57. Deady M, Peters D, Lang H, Calvo R, Glozier N, Christensen H, et al. Designing smartphone mental health applications for emergency service workers. Occup Med (Lond) 2017 Aug 01;67(6):425-428. [doi: 10.1093/occmed/kqx056] [Medline: 28535246]
58. Colder Carras M, Kalbarczyk A, Wells K, Banks J, Kowert R, Gillespie C, et al. Connection, meaning, and distraction: a qualitative study of video game play and mental health recovery in veterans treated for mental and/or behavioral health problems. Soc Sci Med 2018 Nov;216:124-132 [FREE Full text] [doi: 10.1016/j.socscimed.2018.08.044] [Medline: 30257787]
59. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
60. Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, et al. Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. J Med Internet Res 2015 Jul 08;17(7):e166 [FREE Full text] [doi: 10.2196/jmir.4391] [Medline: 26155878]

Abbreviations

ARIA: Alberta Rating Index for Apps

XSL•FO
RenderX

CAF: Canadian Armed Forces
HCP: health care professional
mHealth: mobile health
MM: military member
OSI: occupational stress injury
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PSP: public safety personnel
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial
VAC: Veterans Affairs Canada

Edited by L Buis; submitted 28.06.21; peer-reviewed by A Teles, J Austin; comments to author 01.09.21; revised version received 09.10.21; accepted 19.11.21; published 19.01.22.

Please cite as:
Voth M, Chisholm S, Sollid H, Jones C, Smith-MacDonald L, Brémault-Phillips S
Efficacy, Effectiveness, and Quality of Resilience-Building Mobile Health Apps for Military, Veteran, and Public Safety Personnel Populations: Scoping Literature Review and App Evaluation
JMIR Mhealth Uhealth 2022;10(1):e26453
URL: https://mhealth.jmir.org/2022/1/e26453
doi:10.2196/26453
PMID:35044307

©Melissa Voth, Shannon Chisholm, Hannah Sollid, Chelsea Jones, Lorraine Smith-MacDonald, Suzette Brémault-Phillips. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 19.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Pulse Oximeter App Privacy Policies During COVID-19: Scoping Assessment

Rachele Hendricks-Sturrup[1,2], DHSc, MSc, MA

[1]Future of Privacy Forum, Washington, DC, United States

[2]Duke-Margolis Center for Health Policy, Washington, DC, United States

Corresponding Author:
Rachele Hendricks-Sturrup, DHSc, MSc, MA
Future of Privacy Forum
1400 I St NW
Suite 450
Washington, DC, 20005
United States
Phone: 1 202 768 8950
Email: rachele.hendricks.sturrup@duke.edu

Abstract

Background: Pulse oximeter apps became of interest to consumers during the COVID-19 pandemic, particularly when traditional over-the-counter pulse oximeter devices were in short supply. Yet, no study to date has examined or scoped the state of privacy policies and notices for the top-rated and most downloaded pulse oximeter apps during COVID-19.

Objective: The aim of this study was to examine, through a high-level qualitative assessment, the state and nature of privacy policies for the downloaded and top-rated pulse oximeter apps during the COVID-19 pandemic to (1) compare findings against comparable research involving other mobile health (mHealth) apps and (2) begin discussions on opportunities for future research or investigation.

Methods: During August-October 2020, privacy policies were reviewed for pulse oximeter apps that had either at least 500 downloads (Google Play Store apps only) or a three out of five-star rating (Apple Store apps only). In addition to determining if the apps had an accessible privacy policy, other key privacy policy–related details that were extracted included, but were not limited to, app developer location (country); whether the app was free or required paid use/subscription; whether an ads disclosure was provided on the app's site; the scope of personal data collected; proportionality, fundamental rights, and data protection and privacy issues; and privacy safeguards.

Results: Six pulse oximeter apps met the inclusion criteria and only 33% (n=2) of the six apps had an accessible privacy policy that was specific to the pulse oximeter app feature (vs the app developer's website or at all). Variation was found in both the regulatory nature and data privacy protections offered by pulse oximeter apps, with notable privacy protection limitations and gaps, although each app provided at least some information about the scope of personal data collected upon installing the app.

Conclusions: Pulse oximeter app developers should invest in offering stronger privacy protections for their app users, and should provide more accessible and transparent privacy policies. This is a necessary first step to ensure that the data privacy of mHealth consumers is not exploited during public health emergency situations such as the COVID-19 pandemic, where over-the-counter personal health monitoring devices could be in short supply and patients and consumers may, as a result, turn to mHealth apps to fill such supply gaps. Future research considerations and recommendations are also suggested for mHealth technology and privacy researchers who are interested in examining privacy implications associated with the use of pulse oximeter apps during and after the COVID-19 pandemic.

(JMIR Mhealth Uhealth 2022;10(1):e30361) doi:10.2196/30361

KEYWORDS

COVID-19; pulse oximeters; mobile apps; mHealth; privacy

XSL•FO
RenderX

Introduction

Symptom and health behavior tracking applications or smartphone apps continue to grow in popularity along with government interest and oversight over the privacy practices of such apps [1]. Notably, recent research has raised concerns about the privacy and security of information provided and exchanged via mobile health (mHealth) and wellness apps in general, especially apps that target certain disease or patient populations. For instance, a recent study that examined 29 commercial smartphone apps developed for individuals coping with migraines or headaches (diary and relaxation apps) concluded that the apps shared information with third parties, while also noting that there are few legal protections that protect against the sale or disclosure of app user information to third parties [2]. Another study deployed a semiautomatic app search module to examine the privacy-related information of diabetes-focused apps available via Android, discovering that nearly 60% of the 497 apps surveyed requested permissions that significantly risk user data privacy and that 28.4% of the apps did not house their privacy policies on a website [3]. Several other recent studies discovered similar variation in findings for a variety of broadly available mHealth apps, which are discussed further below [4-15].

Pulse oximeter apps became of interest to consumers during the COVID-19 pandemic, particularly when traditional over-the-counter pulse oximeter devices were in short supply, as consumers sought to personally monitor themselves for hallmark symptoms of SARS-CoV-2 infection (eg, low blood oxygen saturation) [16,17]. Traditional medical-grade pulse oximeters function using a clamp that can be placed over a person's fingertip, which then shines a light over the fingertip to measure blood oxygen saturation. Some pulse oximeter apps connect with traditional, medical-grade pulse oximeter devices via Bluetooth or USB and can export data/records to other devices. The US Food and Drug Administration (FDA) defines a pulse oximeter as "a device used to transmit radiation at a known wavelength(s) through blood and to measure the blood oxygen saturation based on the amount of reflected or scattered radiation" [18]. To obtain a pulse oximetry reading, pulse oximeter devices project a light at a specific wavelength that is shined over a specific area of a person's body while the device measures how much light is absorbed (vs transmitted) by the blood cells within that area of the body. This process is somewhat similar to how mobile apps collect these same measurements, and studies have examined the differences in performance between pulse oximeter mobile apps and medical/hospital-grade pulse oximeters [19,20].

Consumer Reports recently outlined the pros and cons of using pulse oximeter apps, noting a specific app that is available on smart Android phones, called the Pulse Oximeter-Heart Rate Oxygen Monitor App, developed by digiDoc Technologies. This app is meant to be used only for athletic or fitness purposes and not for medical purposes given its technical performance limitations [21]. However, pulse oximeter apps that rely on flash and camera lighting to measure blood oxygen saturation are not always reviewed and approved by regulatory authorities such as the US FDA. Pulse oximeters recently underwent increased scrutiny during the COVID-19 pandemic due to research highlighting racial bias in pulse oximeter devices developed and trained on nonracially diverse populations of individuals, thereby prompting the need for further investigation regarding the scientific validity and accuracy of pulse oximeters [22].

Traditional, over-the-counter pulse oximeters became in short supply during the pandemic amid supply chain shortages. Yet, no study has been published to date broadly examining the privacy policies of pulse oximeter apps at the height of the broad societal impact of the COVID-19 pandemic (mainly, during 2020). Specifically, the literature offers no high-level qualitative assessment on the state or nature of privacy policies for the most downloaded and top-rated pulse oximeter apps during this challenging period. Therefore, the aim of this study was to address this gap to compare findings against comparable research involving other mHealth apps, which can begin discussions on how future research can fill important knowledge gaps about the state of privacy practices for pulse oximeter apps during and after the COVID-19 pandemic.

Methods

In August 2020, the Google Play Store and Apple Store were searched to scope and identify pulse oximeter apps that had either at least 500 downloads (Google Play Store apps only) or a three out of five-star rating (Apple Store apps only). The total number of pulse oximeter apps available on both the Google Play Store and Apple Store was not tallied for purposes of the analysis. Under the direction of the author, two junior analysts reviewed privacy policies for pulse oximetry–specific apps that met the inclusion criteria between August and October 2020.

The following information was extracted from policies and statements found on the app developers' publicly available websites and respective app stores: software purpose; developer location (country); whether the app was free or required paid use/subscription; mobile device access permissions stated on the app's download site; whether an ads disclosure was provided on the app's site; scope of personal data collected; how personal data are collected; who can access personal data; why personal data are used; where the data are stored; how long the data are stored; proportionality, fundamental rights, and data protection and privacy issues; privacy safeguards; and whether the privacy policy was accessible via the app store.

This specific information was extracted to align with our prior work to examine the extent to which each pulse oximeter app "appropriately and ethically balanced public health and safety with privacy risks and other interferences with civil liberties" during the COVID-19 pandemic [23].

These details were captured and summarized independently by the same two junior analysts and the summary was reviewed by RHS for accuracy and clarity. The finalized summary of findings was not reviewed and verified by the developers of the apps that met the inclusion criteria for further accuracy.

XSL•FO
RenderX

Results

Descriptive Assessment

Six apps in total met the study-specific inclusion criteria. Three of these six apps connect to or are compatible with an externally associated oximeter device. Among these three, only one provided a statement of FDA approval as a pulse oximeter device (EMAY Bluetooth Pulse Oximeter). The app developer's headquarter locations were disclosed for all except one of the six apps (OxyCare-[Pulse Oximeter]); apps were developed in Vietnam, Spain, the United States, China, and Canada. Two apps required payment to either download or access certain features within the app (Pulse Oximeter-Beat & Oxygen and Oxxiom).

Privacy Notice Assessment

Table 1 provides a full summary of privacy policy provisions and considerations for each of the six pulse oximeter apps reviewed.

Only two of the apps covered in this review (Pulse Oximeter-Beat & Oxygen and Kenek Edge) had privacy policies that were accessible directly via the app store. The other four apps reviewed (Oximeter, OxyCare [Pulse Oximeter], Oxxiom, and EMAY Bluetooth Pulse Oximeter) either did not have privacy policies that are accessible directly via the app store or did not have an accessible privacy policy that is specific to the pulse oximeter app. However, one app offered a user guide that contains user privacy guidance (Oxxiom). One app's privacy policy is specific to the developer's website versus the pulse oximeter app (EMAY Bluetooth Pulse Oximeter).

All six apps reviewed provided some information about the scope of personal data collected upon installing the app. All but one app (OxyCare [Pulse Oximeter]) specifically described *how* personal data are collected, *who* can access the personal data, *why* personal data are used, and *where and for how long* personal data are stored. Half of the apps reviewed (Pulse Oximeter-Beat & Oxygen, Oximeter, and Kenek Edge) provide an ads disclosure directly on the app download site. Two apps (OxyCare-[Pulse Oximeter] and Oxxiom) did not disclose deidentification commitments within the scope of proportionality, fundamental rights, and data protection and privacy issues. None of the apps' policies explicitly stated if personal data would be used for research purposes. Only one app's policy (Oximeter) explicitly stated that personal data are deleted once the app user permanently deletes the account. The five app developers that described *who* can access personal data in their privacy notices (excluding OxyCare [Pulse Oximeter]) discussed circumstances in which personal data are collected from and used by nonusers (ie, third-party service providers, advertising partners). Two of those five apps explicitly describe personal data access/use by law enforcement (Oximeter and Kenek Edge). Data collection and use for four of the apps are explicitly "opt-in" (Pulse Oximeter-Beat & Oxygen, Oximeter, Oxxiom, and Kenek Edge) and one app explicitly recommends disabling cookies as a privacy safeguard for personal data (EMAY Bluetooth Pulse Oximeter).

XSL•FO
RenderX

Table 1. Summary of pulse oximeter app privacy policy provisions reviewed during August-October 2020.

Category	Pulse Oximeter-Beat & Oxygen	Oximeter	OxyCare-(Pulse Oximeter)	Oxxiom	EMAY Bluetooth Pulse Oximeter	Kenek Edge
Software purpose	General digital health management app that helps users personally check their blood oxygen level and heart rate at any time	General digital health management app that helps users see the percentage of breathable oxygen at their current altitude and check what percentage of oxygen they are breathing	Digital health app that connects to traditional, medical-grade pulse oximeters via Bluetooth or USB	Digital health app that works only with the Oxxiom pulse oximetry system/device	Digital health app that allows users to transfer the pulse oximetry and heart rate data from the EMAY Bluetooth Pulse Oximeter device (Food and Drug Administration–approved) to smartphones	General digital health management app that helps users measure their blood oxygen and heart rate using a hospital-grade finger sensor that can be attached to users' mobile phones or tablets
Developer location (country)	Vietnam	Spain	Not disclosed	United States	China	Canada
Free/Paid	Free to install but charges per feature offered within the app	Free to install and use	Free to install and use	Charge to install; pulse oximeter sold separately	Free to install and use	Free to install and use
Mobile device access permissions stated on app download site	Storage; Wi-Fi connection information; wearable sensors/activity data; photos, media, and files; receive data from internet; full network access; prevent device from sleeping; view network connections; run at startup; control vibration	Location; photos, media, and files; storage; view network connections; full network access	Location; photos, media, and files; storage; pair with Bluetooth devices; access Bluetooth settings	Users may post, upload, store, share, send, or display photos, images, video, data, text, comments, and other information and content ("Your Content") to and via the app, which would grant the app a nonexclusive, transferable, sublicensable, worldwide, royalty-free license to use, copy, modify, publicly display, reproduce, translate, and distribute user content	Not disclosed	Location; weblogs; IP address; web browser information; date and time user accessed or left the developer's website and which pages the user viewed; behavioral data (eg, sleep patterns); user communication records with the developer; personal information (eg, name, age, gender, height, and weight)
Ads disclosure on app download site?	Yes	Yes	No	No	No	Yes
Scope of personal data collected	"Registration" data (eg, name, email); "transaction" data (eg, purchases, offer responses, downloads); "help" data; app use (eg, heart rate, steps, flights climbed, age, height, weight); other data (eg, mobile device type, unique device ID, IP address, mobile operating system, mobile internet browsers)	"Account" data (eg, username, password, email); "additional" data (eg, biography, location, website, picture, address book); location data (eg, mobile or IP address); "log data" (eg, IP address, browser type, operating system, referring webpage, pages visited, location, mobile carrier, device information, search terms, cookies)[a]	Location (approximate via network and precise via GPS); USB storage (photos, media, files)[a]	Date and times of measurements; SpO2[b], PR[c], and PI[d] measurements; sale information (eg, shipping address, contact information, credit card information)	Deidentified "basic" web server visitor information (eg, IP address, browser details, timestamps, referring pages)	Visit data (eg, location data, weblogs and other communication data, IP address, web browser information, date and time accessed); form data (eg, name, email); sleep data (eg, actions, behaviors, treatments, medication, and general wellness); identifying information (eg, email, device ID, site password); personal information (eg, name, age, gender, height, weight); location information

XSL•FO
RenderX

Category	Pulse Oximeter-Beat & Oxygen	Oximeter	OxyCare-(Pulse Oximeter)	Oxxiom	EMAY Bluetooth Pulse Oximeter	Kenek Edge
How personal data are collected	Via individuals (account creation or contacting the app); automatic app collection (eg, device, IP address); and third-party tracking technology (eg, cookies)	Via "various websites, email notifications, apps, buttons, widgets, ads, and commerce services"[a]	Not disclosed	Self-reported and self-uploaded	Tracking via cookies	Via individuals (account creation, contacting the app/site); automatic collection (eg, device, IP address); and third-party tracking technology (eg, cookies)
Who can access personal data	Authorized employees and contractors, service providers, app partners, advertisers, advertising networks. Users can opt-out from third-party use of data by uninstalling the app	If the user decides to publish the information, it will be public: service providers, third-party apps, and websites when the user links accounts, sellers of goods and services, law enforcement[a]	Not disclosed	Third-party payment service providers and authorized third-party e-commerce websites	Advertising partners and other third parties who use cookies	Access via business transfers, law enforcement, and via consent to third parties. Customer PII[e] is not available to third-party advertisers; however, these third parties may automatically collect other information via cookies
Why personal data are used	To contact individuals, advertise relevant products and services, to use the app	To provide the app services while improving them over time and to provide relevant advertising[a]	Not disclosed	To provide app services	For routine administration and maintenance purposes	To contact individuals, advertise via third parties, perform the app's services, and comply with the law
Where the data are stored	Internal memory of the user's cellular device. Data processing takes place in the United States	Internal memory of the user's device(s). Data processing takes place in the United States and any country where the app operates[a]	Not disclosed	Internal memory of the user's iOS device	Not disclosed	Internal memory of the user's devices; otherwise, not disclosed
How long the data are stored	Data for advertising purposes are stored as long as the app is installed on the mobile phone	If the user permanently deletes the account, then the data are deleted. Log data are deleted after a few months[a]	Not disclosed	Credit card information is not stored	Not disclosed	Not disclosed
Proportionality, fundamental rights, and data protection and privacy issues	Only aggregated, anonymous data are "periodically" transmitted to third parties. Advertisers will only have access to "Automatically Collected Information," which is the device's unique ID, IP address, mobile operating system, type of mobile browsers, and app use information	Nonprivate, aggregated, or "otherwise nonpersonal information" will be shared or disclosed[a]	Not disclosed	Not disclosed	User's personal information cannot be used to identify specific visitors	Individuals can visit the app/website without revealing any personal information

XSL•FO
RenderX

Category	Pulse Oximeter-Beat & Oxygen	Oximeter	OxyCare-(Pulse Oximeter)	Oxxiom	EMAY Bluetooth Pulse Oximeter	Kenek Edge
Privacy safeguards	The app is opt-in. Physical, electronic, and procedural safeguards of data (eg, authorization process)	The app is opt-in[a]	Not disclosed	The app is opt-in	The app recommends disabling cookies	The app is opt-in. The developer has a "commercially suitable physical, electronic, and managerial procedure" to safeguard and secure collected information
Privacy policy accessible via app store? [24]	Yes [25]	No (same app developer's Privacy Policy: RamLabs) [26][a]	No	No [27]	Yes, although specific to the company website versus the app [28]	Yes [29]

[a]Information taken from the app developers' general privacy policies; the policy could apply to the pulse oximeter app reviewed or a different app made by the developer.

[b]SpO2: oxygen saturation.

[c]PR: pulse rate.

[d]PI: perfusion index.

[e]PII: personal identifiable information.

Discussion

Principal Findings

The present findings fill an important literature gap regarding the privacy policies of pulse oximeter apps during the COVID-19 pandemic. These findings are largely consistent with trends observed in prior research that has examined the accessibility, structure, and substance of commercial mHealth apps' privacy policies [2-15]. Namely, the top-rated or the most downloaded pulse oximeter apps during the COVID-19 pandemic either did not provide accessible privacy policies via the app store or did not provide privacy policies that were specific to the pulse oximeter app being offered. Thus, the present findings seemingly align with observations seen in recent assessments of privacy policies for a variety of mHealth apps. Although each pulse oximeter app provided some information to users about their scope of data collection, what is perhaps most concerning from a privacy standpoint is that all but one app (OxyCare [Pulse Oximeter]) provided privacy disclosures that are consistent with current privacy recommendations and best practices as well as policy-based guidance.

Limitations

There are limitations to the present analysis and findings such that the observations reported herein are limited to only the highest rated or most downloaded pulse oximeter apps, which effectively excludes pulse oximeter apps that have lower ratings or are downloaded less frequently. In addition, this analysis did not include technical verification and quality assessment criteria for the apps, such as pulse oximeter app usability. Within these limitations are opportunities for further research to explore these important components as a critical next step to this broad analysis. This study was also cross-sectional in time such that it was intentionally limited to capture the state of pulse oximeter app privacy policies at the height of the COVID-19 pandemic when traditional, over-the-counter pulse oximeters were in short supply. Future research should examine if and the extent to which popular pulse oximeter app privacy policies have been either developed or updated.

Alignment With Prior Research Examining the Privacy Policies of mHealth Apps

Several recent studies examined privacy policies and notices for a wide range of mHealth apps, noting trends that are similar to those found in the present analysis of pulse oximeter apps during COVID-19 [2-15]. The Future of Privacy Forum also published a similar study in a 2016 white paper, where they examined whether the most popular free and paid mHealth apps "provided users with access to a privacy policy, and whether the privacy policy was linked from the app's listing page on the iOS [Apple] and Android app marketplaces" [30]. Therefore, the present analysis offers an opportunity to understand how the overall accessibility of privacy policies and notices for pulse oximeter apps during the COVID-19 pandemic compare with that of other health apps generally based on findings from comparable work published within the past 5 years (see Table 2).

XSL•FO
RenderX

Table 2. Comparison of present findings against comparable and prior privacy policy research focused on mobile health (mHealth) apps.

Study	mHealth app category surveyed	Apps meeting inclusion criteria, N	mHealth apps surveyed with an accessible mHealth app privacy policy, n (%)
This study	Pulse oximeter apps during COVID-19	6	2 (33)
FPF Mobile Apps Study [30]	Health and fitness, period tracking, sleep aid	25	19 (76)
Flors-Sidro et al [3]	Diabetes	497	139 (28.0)
O'Loughlin et al [5]	Depression	116	57 (49.1)
Grindrod et al [6]	Medication use and management	185	63 (34.1)
Rosenfeld et al [8]	Dementia	72	33 (46)
Huckvale et al [12]	Mental health (depression and smoking cessation)	36	25 (69)
Bachiri et al [15]	Pregnancy monitoring	38	18 (47)

The findings of this study showed a relatively low percentage of the most downloaded or top-rated pulse oximeter apps during COVID-19 that provided an accessible privacy policy (33%) compared with the average for the current trend seen in the literature for various mHealth apps (50%). This is problematic given that pulse oximeter apps grew in popularity during the COVID-19 pandemic, leaving pulse oximeter app users with an overall low degree of certainty about the privacy and security of their personal data that could be collected, shared, or processed by or via the apps.

Future Opportunities and Priorities for Privacy Researchers and App Developers

Based on the present findings, it is recommended that future privacy research on pulse oximeter apps involve a deeper comparative analysis that would investigate the effectiveness of available privacy policies and/or offer a more technical analysis of privacy and security implications. Future work might also involve a systematic review, meta-analysis, or meta-synthesis of mHealth apps to more robustly capture and compare the state and substance of privacy policies and notices for mHealth apps, including pulse oximeter apps. Moreover, given that (1) certain pulse oximeter app user data could be considered as sensitive data under the EU General Data Protection Regulation (GDPR), and (2) each of these apps could function within the European Union and must therefore comply with the EU GDPR, future work should involve a robust risk assessment of pulse oximeter app and other mHealth app privacy policies against specific articles within the EU GDPR, most notably articles focused on user informed consent, data minimization, legal basis or grounds for data collection, data subjects' rights, and consequential areas [31]. Lastly, pulse oximeter app developers should clarify within their privacy policies their purpose and need to collect sensitive information (eg, geolocation data, browsing data, address book data), as it may be unclear or not intuitive among users why the pulse oximeter app would need to collect such data to provide its intended services or experience to its users, and thus may be perceived as privacy-invasive.

Conclusion

It is clear from the present review and related literature that mHealth apps, including pulse oximeter apps, hold vast opportunities—and perhaps necessity during and after the COVID-19 pandemic—to make their privacy policies more robust and aligned with these current privacy best practices and regulatory requirements. As the practice of medicine becomes increasingly digitized, offering consumers greater options to self-engage in health monitoring and data reporting using personal smartphones, the privacy and security of person-generated health data and traditional health become tantamount. Robust mHealth app consumer or user privacy protections, including, but not limited to, having an accessible and transparent privacy policy, are therefore needed to ensure that the data privacy of mHealth consumers cannot become exploited during public health emergency situations such as the COVID-19 pandemic, if patients and consumers feel compelled to purchase and download mHealth apps in response to short supplies of more traditional, over-the-counter personal health monitoring devices.

Acknowledgments

Gratitude is extended to Pollyanna Sanderson, Policy Counsel, and Veronica Alix, former policy intern, at the Future of Privacy Forum in Washington, DC, for their time and effort put forth to conduct the privacy notice and policy analysis reported herein. RHS is the former Health Policy Counsel and Lead at the Future of Privacy Forum in Washington, DC. The views herein do not necessarily reflect those of Future of Privacy Forum supporters or board members.

Conflicts of Interest

RHS is presently employed by the Duke-Margolis Center for Health Policy and reports contract work with the National Alliance Against Disparities in Patient Health.

XSL•FO
RenderX

References

1. Plant M. Does your health app protect your sensitive info? Federal Trade Commission Consumer Information. 2021 Jan 13. URL: https://www.consumer.ftc.gov/blog/2021/01/does-your-health-app-protect-your-sensitive-info [accessed 2021-08-04]
2. Minen MT, Stieglitz EJ, Sciortino R, Torous J. Privacy issues in smartphone applications: an analysis of headache/migraine applications. Headache 2018 Jul 04;58(7):1014-1027 [FREE Full text] [doi: 10.1111/head.13341] [Medline: 29974470]
3. Flors-Sidro JJ, Househ M, Abd-Alrazaq A, Vidal-Alaball J, Fernandez-Luque L, Sanchez-Bocanegra CL. Analysis of diabetes apps to assess privacy-related permissions: systematic search of apps. JMIR Diabetes 2021 Jan 13;6(1):e16146 [FREE Full text] [doi: 10.2196/16146] [Medline: 33439129]
4. Powell AC, Singh P, Torous J. The complexity of mental health app privacy policies: a potential barrier to privacy. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e158 [FREE Full text] [doi: 10.2196/mhealth.9871] [Medline: 30061090]
5. O'Loughlin K, Neary M, Adkins EC, Schueller SM. Reviewing the data security and privacy policies of mobile apps for depression. Internet Interv 2019 Mar;15:110-115 [FREE Full text] [doi: 10.1016/j.invent.2018.12.001] [Medline: 30792962]
6. Grindrod K, Boersema J, Waked K, Smith V, Yang J, Gebotys C. Locking it down: The privacy and security of mobile medication apps. Can Pharm J (Ott) 2017;150(1):60-66 [FREE Full text] [doi: 10.1177/1715163516680226] [Medline: 28286594]
7. Benjumea J, Ropero J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Assessment of the fairness of privacy policies of mobile health apps: scale development and evaluation in cancer apps. JMIR Mhealth Uhealth 2020 Jul 28;8(7):e17134 [FREE Full text] [doi: 10.2196/17134] [Medline: 32720913]
8. Rosenfeld L, Torous J, Vahia IV. Data security and privacy in apps for dementia: an analysis of existing privacy policies. Am J Geriatr Psychiatry 2017 Aug;25(8):873-877. [doi: 10.1016/j.jagp.2017.04.009] [Medline: 28645535]
9. Zapata B, Niñirola A, Fernández-Alemán J, Toval A. Assessing the privacy policies in mobile personal health records. 2014 Presented at: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2014; Chicago, Illinois p. 4956-4959. [doi: 10.1109/embc.2014.6944736]
10. Benjumea J, Ropero J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Privacy assessment in mobile health apps: scoping review. JMIR Mhealth Uhealth 2020 Jul 02;8(7):e18868 [FREE Full text] [doi: 10.2196/18868] [Medline: 32459640]
11. Papageorgiou A, Strigkos M, Politou E, Alepis E, Solanas A, Patsakis C. Security and privacy analysis of mobile health applications: the alarming state of practice. IEEE Access 2018;6:9390-9403. [doi: 10.1109/ACCESS.2018.2799522]
12. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. JAMA Netw Open 2019 Apr 05;2(4):e192542 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2542] [Medline: 31002321]
13. Hutton L, Price BA, Kelly R, McCormick C, Bandara AK, Hatzakis T, et al. Assessing the privacy of mHealth apps for self-tracking: heuristic evaluation approach. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e185 [FREE Full text] [doi: 10.2196/mhealth.9217] [Medline: 30348623]
14. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
15. Bachiri M, Idri A, Fernández-Alemán JL, Toval A. Evaluating the privacy policies of mobile personal health records for pregnancy monitoring. J Med Syst 2018 Jun 29;42(8):144 [FREE Full text] [doi: 10.1007/s10916-018-1002-x] [Medline: 29959535]
16. Khalid A. Pulse oximeters are selling out because of the pandemic. Quartz. URL: https://qz.com/1832464/pulse-oximeters-for-coronavirus-unnecessary-but-selling-strong/ [accessed 2021-04-28]
17. Singh HJL, Couch D, Yap K. Mobile health apps that help with COVID-19 management: scoping review. JMIR Nurs 2020 Aug 6;3(1):e20596 [FREE Full text] [doi: 10.2196/20596] [Medline: 32897271]
18. Pulse Oximeters - Premarket Notification Submissions 510(k)s: Guidance for Industry and Food and Drug Administration Staff. Food and Drug Administration. 2013. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pulse-oximeters-premarket-notification-submissions-510ks-guidance-industry-and-food-and-drug [accessed 2021-08-04]
19. Browne SH, Bernstein M, Pan SC, Gonzalez Garcia J, Easson CA, Huang C, et al. Smartphone biosensor with app meets FDA/ISO standards for clinical pulse oximetry and can be reliably used by a wide range of patients. Chest 2021 Feb;159(2):724-732 [FREE Full text] [doi: 10.1016/j.chest.2020.08.2104] [Medline: 32926871]
20. Tarassenko L, Greenhalgh T. Question: Should smartphone apps be used clinically as oximeters? Answer: No. The Centre for Evidence-Based Medicine. 2020 Apr 01. URL: https://www.cebm.net/covid-19/question-should-smartphone-apps-be-used-as-oximeters-answer-no/ [accessed 2021-08-04]
21. Peachman R. People Concerned About COVID-19 Are Using Pulse Oximeters to Measure Oxygen Levels. These Are the Pros and Cons. People Concerned About COVID-19 Are Using Pulse Oximeters to Measure Oxygen Levels. URL: https://www.consumerreports.org/medical-symptoms/covid-19-pulse-oximeters-oxygen-levels-faq/ [accessed 2021-04-28]
22. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. N Engl J Med 2020 Dec 17;383(25):2477-2478 [FREE Full text] [doi: 10.1056/NEJMc2029240] [Medline: 33326721]
23. FPF charts the role of mobile apps in pandemic response. Future of Privacy Forum. 2020 Apr 03. URL: https://fpf.org/blog/fpf-charts-the-role-of-mobile-apps-in-pandemic-response-chart/ [accessed 2021-11-11]

XSL•FO
RenderX

24. OxyCare - (pulse oximeter) - apps on Google Play. Google. URL: https://play.google.com/store/apps/details?id=com.berry_med.oxycare&hl=en_US&gl=US [accessed 2021-04-28]
25. Privacy policy. PVDAPPS. URL: http://privacy.pvdapps.com/privacy.html [accessed 2021-04-28]
26. World Summits privacy policy. WorldSummits. URL: http://worldsummits.net/Legal/Privacy_en-EN.aspx [accessed 2021-04-28]
27. oxxiom user guide. true wearables. URL: https://www.truewearables.com/img/UserGuide.pdf [accessed 2021-04-28]
28. Privacy policy. EMAY. URL: https://www.emayltd.com/http-www-emayltd-com-index-php-Home-Index-wm-id-7.html [accessed 2021-04-28]
29. Privacy statement. Kenek Health Support Portal. URL: https://kenekhealth.freshdesk.com/support/solutions/articles/47000272334-privacy-statement [accessed 2021-04-28]
30. FPF mobile apps study. Future of Privacy Forum. 2016 Aug. URL: https://fpf.org/wp-content/uploads/2016/08/2016-FPF-Mobile-Apps-Study_final.pdf [accessed 2021-11-11]
31. Wolford B. Writing a GDPR-compliant privacy notice (template included). GDPR. URL: https://gdpr.eu/privacy-notice/ [accessed 2021-08-04]

Abbreviations

FDA: Food and Drug Administration
GDPR: General Data Protection Regulation
mHealth: mobile health

Edited by L Buis; submitted 11.05.21; peer-reviewed by D Mendelsohn, I Schiering, J Ropero; comments to author 02.07.21; revised version received 05.08.21; accepted 03.12.21; published 27.01.22.

Please cite as:
Hendricks-Sturrup R
Pulse Oximeter App Privacy Policies During COVID-19: Scoping Assessment
JMIR Mhealth Uhealth 2022;10(1):e30361
URL: https://mhealth.jmir.org/2022/1/e30361
doi:10.2196/30361
PMID:35084348

©Rachele Hendricks-Sturrup. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 27.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Smartphone App to Improve Oral Anticoagulation Adherence in Patients With Atrial Fibrillation: Prospective Observational Study

Keitaro Senoo[1,2], MD; Tomonori Miki[2], MD, PhD; Takashi Ohkura[2], MD; Hibiki Iwakoshi[2], MD; Tetsuro Nishimura[2], MD; Hirokazu Shiraishi[1,2], MD, PhD; Satoshi Teramukai[3], PhD; Satoaki Matoba[1,2], MD, PhD

[1]Department of Cardiac Arrhythmia Research and Innovation, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

[2]Department of Cardiovascular Medicine, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

[3]Departments of Biostatistics, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

Corresponding Author:
Keitaro Senoo, MD
Department of Cardiac Arrhythmia Research and Innovation
Graduate School of Medical Science, Kyoto Prefectural University of Medicine
Kyoto-shi, Kamigyo-ku Kajii-cho 465, Kawaramachi-Hirokoji
Kyoto, 602-8566
Japan
Phone: 81 8031117168
Email: swcqg251@yahoo.co.jp

Abstract

Background: Poor adherence to oral anticoagulation in elderly patients with atrial fibrillation (AF) has been shown to negatively impact health care costs, morbidity, and mortality. Although various methods such as automated reminders, counseling, telephone support, and patient education have been effective in improving medication adherence, the burden on health care providers has been considerable. Recently, an attempt has been made to improve medication adherence without burdening health care providers by using smartphone apps; however, the use of the app for elderly patients with AF is still limited.

Objective: The purpose of this study was to determine whether the newly developed smartphone app for patients with AF (the Smart AF), which integrates education, automatic reminder, and patient engagement strategies with a simple user interface, can improve medication adherence in elderly patients with AF.

Methods: Patient enrollment was carried out by obtaining informed consent from patients with AF attending Kyoto Prefectural University of Medicine hospital between May 2019 and September 2020. Follow-up was planned at 1, 3, and 6 months after enrollment, and questionnaire reminders were automatically sent to patient apps at designated follow-up time points. A questionnaire-based survey of medication adherence was performed electronically using the self-reported 8-item Morisky Medication Adherence Scale (MMAS-8) as the survey tool.

Results: A total of 136 patients with AF were enrolled in this study. During the follow-up period, 112 (82%) patients underwent follow-up at 1 month, 107 (79%) at 3 months, and 96 (71%) at 6 months. The mean age of the enrolled patients was 64.3 years (SD 9.6), and male participants accounted for 79.4% (108/136) of the study population. The mean CHADS2 (congestive heart failure, hypertension, age, diabetes, previous stroke, or transient ischemic attack) score was 1.2, with hypertension being the most common comorbidity. At the time of enrollment, 126 (93%) and 10 (7%) patients were taking direct oral anticoagulants and warfarin, respectively. For medication adherence as measured according to the MMAS-8, MMAS scores at 1 month, 3 months, and 6 months were significantly improved compared with baseline MMAS scores (all P values less than .01). The overall improvement in medication adherence achieved by the 6-month intervention was as follows: 77.8% (14/18) of the patients in the high adherence group (score=8) at baseline remained in the same state, 45.3% (24/53) of the patients in the medium adherence group (score=6 to <8) at baseline moved to the high adherence group, and 72% (18/25) of the patients in the low adherence group (score <6) moved to either the medium or high adherence group.

Conclusions: The Smart AF app improved medication adherence among elderly patients with AF. In the realm of medication management, an approach using a mobile health technology that emphasizes education, automatic reminder, and patient engagement may be helpful.

(JMIR Mhealth Uhealth 2022;10(1):e30807) doi:10.2196/30807

XSL•FO
RenderX

KEYWORDS

atrial fibrillation; smartphone app; anticoagulants; drug adherence; education; patient involvement

Introduction

Atrial fibrillation (AF), the most common chronic arrhythmia, affected 33.5 million people worldwide, in 2010, and the number is projected to double by 2030 [1,2]. AF is associated with a 5-fold increased risk of stroke, and oral anticoagulation (OAC) is required for people at moderate-to-high risk of stroke [3]. Poor adherence to medication in the real world can alter efficacy and safety estimates from randomized controlled trials, leading to poorer health outcomes and greater health care costs [4,5]. A recent meta-analysis found that suboptimal adherence to and persistence with direct oral anticoagulants (DOACs) is common [6]. For example, patients with AF do not take a DOAC every 4 days; one-third of the patients have less than 80% adherence; real-world persistence with DOACs is lower than in randomized controlled trials; and patients with poor adherence have a higher risk of stroke.

To date, various greater efforts in monitoring and interventions have been used to improve OAC adherence. For instance, the AEGEAN (Assessment of an Education and Guidance Programme for Eliquis Adherence in Non-Valvular Atrial Fibrillation) trial explored the impact of education (ie, using booklets and reminder tools) and telephone follow-up using a virtual clinic on adherence to apixaban. However, in the AEGEAN trial, electronically measured adherence did not differ between the usual care and intervention groups, with adherence rates of 88.5% and 88.3%, respectively [7]. A study by Shore et al [8] reported that enhanced pharmacist engagement and longer patient monitoring and follow-up were associated with greater adherence to dabigatran. In a study by Desteghe et al [9], remote monitoring of daily medication uses and daily remote monitoring and individualized feedback by telephone resulted in very high adherence to DOACs, with 99.0% demonstrating adherence and 96.8% demonstrating regimen adherence. FACILITA (strategies for improving dabigatran adherence for stroke prevention in patients with non-valvular atrial fibrillation) study revealed that a mixed intervention, consisting of patient education and a simple calendar reminder for drug intake, was an effective strategy to improve adherence to dabigatran (to 91% and 89% at 6 and 12 months, respectively, compared with 65% and 63% for the control group) [10]. To improve medication adherence, various methods such as reminders, counseling, telephone support, and patient education are effective; however, these long-term interventions impose a considerable burden on health care providers [11,12].

Recently, attempts have been made to increase medication adherence without burdening health care providers by using smartphone apps in various fields [13-15]. The Health Buddies app (DAE Studios) was developed as a tool to improve adherence by providing a virtual contract with the patients' grandchildren [16]. The mAF app was developed to integrate clinical decision support, education, and patient-involvement strategies [17]. Although some success was achieved in these studies, there were problems such as the complexity of the app user interface.

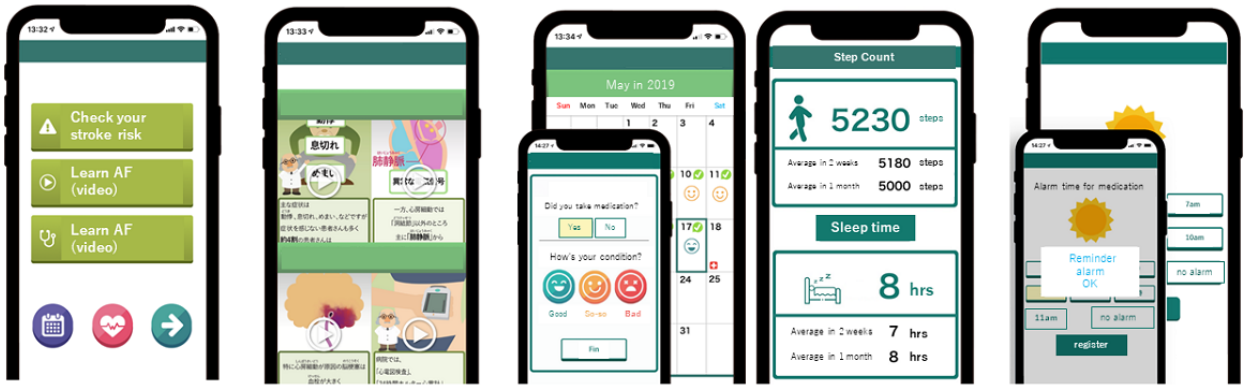
The purpose of this study was to determine whether the newly developed smartphone app for patients with AF (the Smart AF app), which integrates education, automatic reminder, and patient engagement strategies with a simple user interface, can improve medication adherence in elderly patients with AF.

Methods

Features of the Smart AF App

The Smart AF app was developed by Health Tech Innovation Center, in association with Kyoto Prefectural University of Medicine, funded by the BMS/Pfizer Japan Thrombosis Investigator Initiated Research Program (JRISTA). There are 3 features of the Smart AF app (Figure 1).

Figure 1. Features of the smartphone app for atrial fibrillation (the Smart AF app).



Educational Program

There are 7 components to the patient educational program. With videos that are approximately 1-2 minutes in length, patients can learn about AF and learn self-management methods, including how to detect and treat AF, the importance of anticoagulation, and the treatment of comorbidities.

Patient Engagement

After inputting information about the characteristics of their AF history, other medical history, medication information (eg,

XSL•FO
RenderX

type of antithrombotic medication or other concomitant medications), and lifestyle behaviors, each patient's CHADS2 score is automatically calculated, enabling physicians to easily understand their risk for stroke. Furthermore, the app can take inputs on step counts, sleep time, and presence or absence of symptoms and, by sharing this information with health care providers, the proactive involvement of the patient in their own care can be promoted and supported.

Reminder Alarm

A reminder is automatically transmitted through the app in the morning and evening daily to prevent forgetting to take medication.

Outcome Measures

Follow-up was planned at 1, 3, and 6 months after enrollment, and a reminder email for the survey was automatically sent to the patient's app at the time of follow-up. The self-reported 8-item Morisky Medication Adherence Scale (MMAS-8) was used as the survey instrument [18-20]. The MMAS-8 score assesses patients' self-reported adherence to their anticoagulant medication. According to the MMAS-8 score (range 0-8), adherence was defined as high (score 8), medium (score 6 to <8) or low (score <6).

Eligibility Criteria for Participants

Patient enrollment was carried out by obtaining informed consent from patients with AF attending Kyoto Prefectural University of Medicine hospitals between May 2019 and September 2020. All participants provided electronic informed consent and were assigned a password. Inclusion criteria were as follows: documented diagnosis of AF; current prescription for OACs (ie, dabigatran, rivaroxaban, apixaban, edoxaban, and warfarin) for at least 3 weeks; and ownership of a mobile phone. Individuals less than 20 years of age, those with valvular AF, and those who had been taking OACs for less than 3 weeks were excluded. This single-center prospective observational study was approved by the Institutional Review Board of the Kyoto Prefectural University of Medicine (ERB-C-1429).

Data Analysis

Statistical analyses were performed using R version 3.6.1 (R Foundation for Statistical Computing). Continuous variables are expressed as mean and SD, and categorical variables as number and percentage. MMAS-8 scores at 4 time points (baseline, 1, 3, and 6 months later) were plotted to illustrate changes in these variables over time. The Wilcoxon signed-rank test was used to evaluate differences in MMAS-8 scores between baseline and 1, 3, and 6 months. For the 96 patients who were able to complete the follow-up up to 6 months, the change in adherence from the time of enrollment to 6 months is reflected graphically by a low, middle, and high MMAS score. The percentage of patients who activated the app at least once per day is expressed as the mobile app retention rate, and the mean retention rate with SD per week at 1, 3, and 6 months was calculated and graphed. Associated factors were also identified using simple regression analysis. "Low retention," defined as the percentage of days in which the app was activated at least once, was ≤10% during the observation period. Differences with $P<.05$ were considered to be statistically significant.

This study was funded by JRISTA. Permission for use of the MMAS-8 scale and its coding has been acquired, and a license agreement is available from MMAR, LLC, Donald E Morisky, ScD, ScM, MSPH, 294 Lindura Ct., United States.

Results

Between May 2019 and September 2020, a total of 136 patients with AF were enrolled in this study. During the follow-up period, 112 (82%) patients underwent the follow-up survey at 1 month, 107 (79%) at 3 months, and 96 (71%) at 6 months. The mean age of the enrolled patients was 64.3 years (SD 9.6), and males accounted for 79.4% (108/136) of the study population; 89.7% (122/136) of the patients were married and had family members currently residing with them; 63.9% (87/136) of the patients had attended college as their highest level of education, and 57.3% (78/136) were currently working; 75% (102/136) were currently practicing regular dietary habits, and 19.8% (27/136) engaged in daily exercise; 74.2% (101/136) of the patients had a history of AF treatment for ≥1 year, and hypertension was the most common comorbidity, with a mean CHADS2 score of 1.2. Moreover, 38.9% (53/136) of the patients experienced palpitation symptoms, 130 (96%) patients with a Hospital Anxiety and Depression Scale (HADS)-A score of ≥8 (major anxiety), 30 (22%) patients with a HAD-D score of ≥11 (major depression), and 95 (70%) patients with a HADS-T score of ≥20 (major anxiety and depression). At enrollment, 126 (93%) patients were taking DOAC, and 10 (7%) were taking oral warfarin (Table 1).

For medication adherence, as measured according to the MMAS-8, the low adherence group was 27.9% (n=38) at baseline, 16.0% (n=18) at 1 month, 15.9% (n=17) at 3 months, and 12.5% (n=12) at 6 months. Medium adherence was 55.1% (n=75) at baseline, 43.8% (n=49) at 1 month, 44.9% (n=48) at 3 months, and 44.8% (n=43) at 6 months. High adherence group was 16.9% (n=23) at baseline, 40.2% (n=45) at 1 month, 39.3% (n=42) at 3 months, and 42.7% (n=41) at 6 months, respectively. Compared with baseline MMAS scores, MMAS scores at 1 month, 3 months, and 6 months were significantly improved (all P values<.01) (Figure 2).

Furthermore, the overall improvement in medication adherence achieved by the intervention was as follows: 77.8% (14/18) of the patients in the high adherence group at baseline remained there; 45.3% (24/53) of the patients in the medium adherence group at baseline moved to the high adherence group; and 72% (18/25) of the patients in the low adherence group moved to either the medium or high adherence groups (Figure 3).

The mobile app retention rate is a plot of the percentage of patients who activate the app at least once per day (Figure 4).

XSL•FO
RenderX

Table 1. Demographics of the participants (N=136).

Characteristics	Values
Age (years), mean (SD)	64.3 (9.6)
Male, n (%)	108 (79.4)
Married, n (%)	122 (89.7)
Living together, n (%)	123 (90.4)
University education level, n (%)	87 (64.0)
Full-time or part-time employed	78 (57.4)
Regular meal, n (%)	102 (75.0)
Habit of exercise, n (%)	
Every day	27 (19.9)
Sometimes	54 (39.7)
None	55 (40.4)
Duration of atrial fibrillation (>1year)	101 (74.3)
Type of oral anticoagulants, n (%)	
DOAC[a], OD[b] (edoxaban, rivaroxaban)	86 (63.2)
DOAC, BID[c] (apixaban, dabigatran)	40 (29.4)
Warfarin	10 (7.4)
CHADS2[d] score (SD)	1.2 (1.1)
Congestive heart failure, n (%)	18 (16.5)
Hypertension, n (%)	65 (59.6)
Age ≥75 (years), n (%)	13 (9.6)
Diabetes mellitus, n (%)	19 (17.4)
Prior stroke or TIA[e], n (%)	9 (8.3)
With symptom, n (%)	53 (39.0)
HADS[f] scale, n (%)	
HADS-A (≥8—major anxiety; total 21 points)	130 (95.6)
HADS-D (≥11—major depression; total 21 points)	30 (22.1)
HADS-T (≥20—major anxiety and depression: total 42 points)	95 (69.9)

[a]DOAC: direct oral anticoagulants.
[b]OD: once a day.
[c]BID: twice a day.
[d]CHADS2: congestive heart failure, hypertension, age, diabetes, previous stroke, or transient ischemic attack.
[e]TIA: transient ischemic attack.
[f]HADS: Hospital Anxiety and Depression Scale.

XSL•FO
RenderX

Figure 2. Morisky Medication Adherence Scale at baseline, 1 month, 3 months, and 6 months.

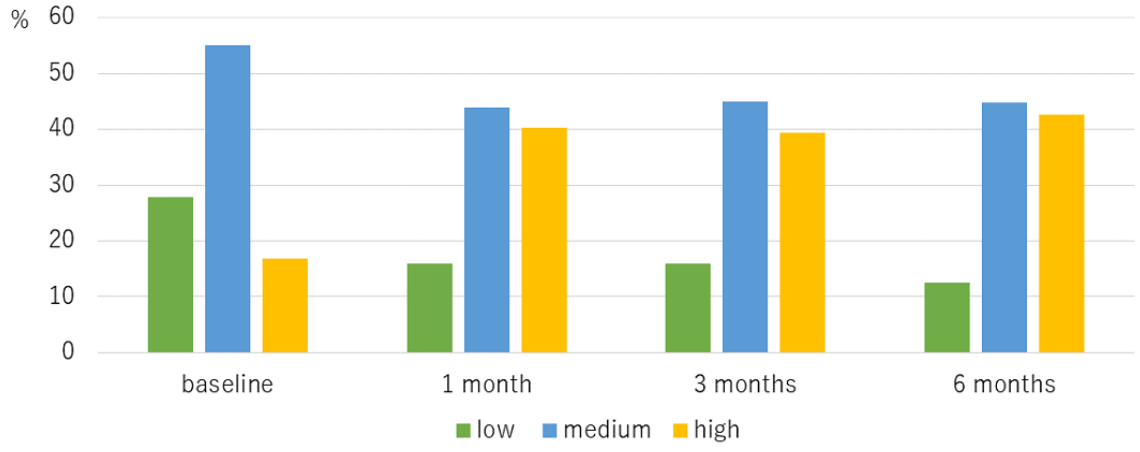


Figure 3. Change in adherence category from baseline to 6 months of follow-up (n=96).

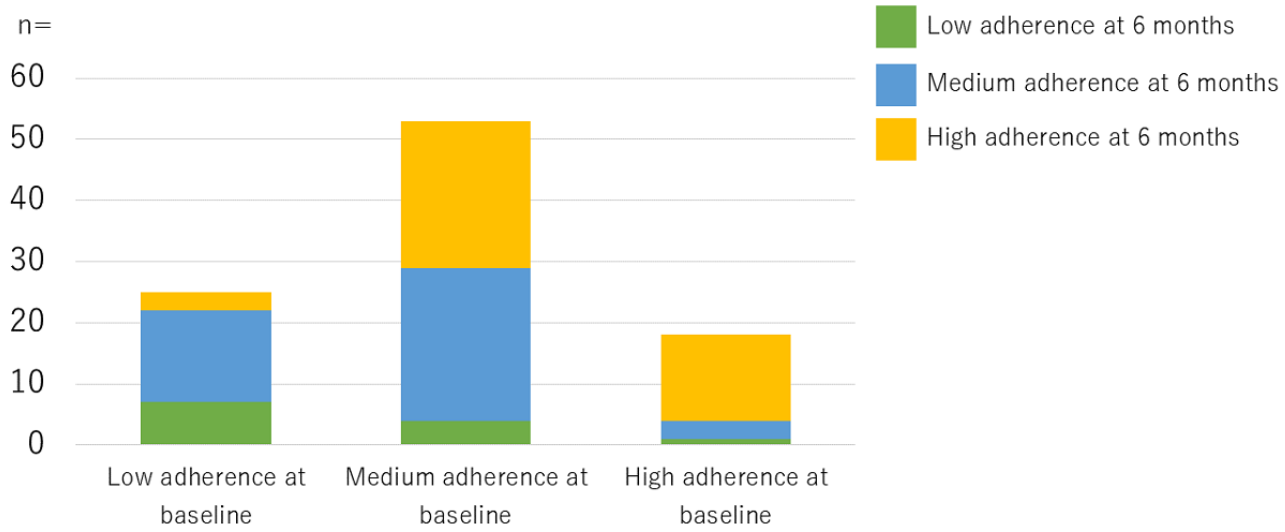
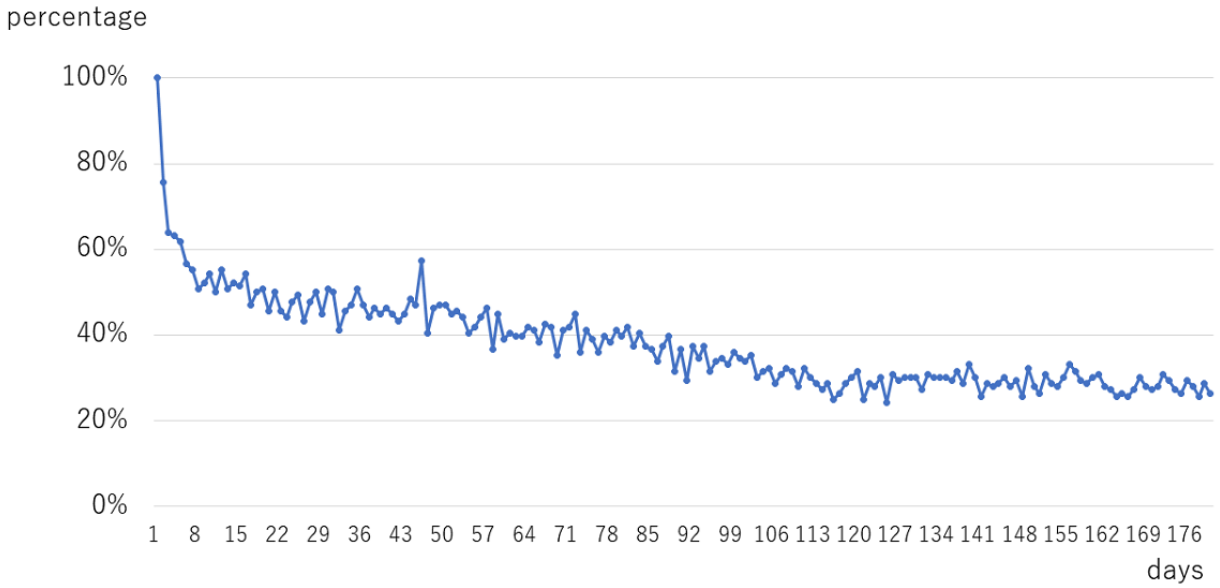


Figure 4. Mobile app retention.



Retention rates for 1, 3, and 6 months were 0.48 (SD 3), 0.36 (SD 3), and 0.27 (SD 1), respectively. In the univariate regression analysis, predictors for a higher retention rate were older age (P=.047), regular meals (P=.04), and HADS-A ≥8 (major anxiety, P=.04), respectively (Table 2).

XSL•FO
RenderX

Table 2. Predictors for retention rate.

Characteristics and variables	Values	
	B[a]	*P* value
Age	0.605	.047
Male	-0.442	.95
Married	-4.134	.67
Living together	-0.349	.97
Education level	-1.756	.77
Employed	0.096	.99
Regular meal	13.56	.04
Habit of exercise		
Every day	3.023	.68
Sometimes	3.871	.52
None	-5.845	.33
Duration of atrial fibrillation (>1 year)	-2.087	.76
Oral anticoagulants, BID[b] (reference as OD[c])	1.538	.81
CHADS2[d] score	-0.257	.90
Polypharmacy	-0.376	.98
With symptom	-1.533	.80
HADS[e] scale		
HADS-A (≥8—major anxiety)	28.875	.04
HADS-D (≥11—major depression)	-10.946	.12
HADS-T (≥20—major anxiety and depression)	-2.533	.69

[a]B: univariate regression analysis.

[b]BID: twice a day.

[c]OD: once a day.

[d]CHADS2: congestive heart failure, hypertension, age, diabetes, previous stroke, or transient ischemic attack.

[e]HADS: Hospital Anxiety and Depression Scale.

Discussion

Main Findings

The Smart AF app has been developed, which integrates patient education, reminder alarm, and patient engagement strategies without increasing the burden on health care providers. Elderly patients with AF who used the Smart AF app demonstrated significantly improved MMAS-8 at 6 months compared to baseline. The Smart AF app demonstrated a significant decrease in the low adherence group and an increase in the high adherence group at 6 months compared with baseline.

Improvement of Adherence

Current clinical guidelines for AF advocate incorporating patient preferences regarding treatment and support and involving patients in management decisions [21,22]. The participation of patients with AF in the process of developing and refining the app through patient involvement resulted in the Smart AF app. For example, to facilitate the use of the app, age-related aspects considered to be useful in the design of mHealth (mobile health) tools, including large screens, large fonts, and ease of navigation, were incorporated into it. As a result, we adopted a simple layout and large navigation buttons that are easy to use, even by elderly individuals. Among the few studies that evaluated the effectiveness of smartphone apps in improving medication adherence in those with AF, research investigating the mAF app revealed that it had a complex design with a great deal of content, such as education programs, clinical decision support material, patient involvement in self-care, structured follow-up, and many areas that users needed to manipulate manually [17]. Regarding the Health Buddies App, there are some problems, such as a lack of interest in the characteristic games of the app, to the point that grandchildren who become “Buddies” must also participate; moreover, although achieving tasks and goals inside the app was fun, long-term behavioral change was difficult because these achievements do not correlate directly with patient health conditions [16].

The Smart AF app was developed with a focus on patient education, reminder alarms, and patient engagement in self-care, and we tried to keep the app simple without any other features.

XSL•FO
RenderX

In fact, the FACILITA study also revealed that a mixed intervention, consisting of patient education and a simple calendar reminder of drug intake, was an effective strategy to improve adherence to dabigatran (to 91% and 89% at 6 and 12 months, respectively, compared with 65% and 63% for the control group) [10]. Therefore, despite the simplicity of the app's contents, we believe that not only did the reminder alarms encourage patients to take their daily medication, but by having them record their medication in the app's calendar, we could encourage their participation and make them more aware of their engagement in self-care by managing their daily health.

Previous studies have also reported that tailor-made educational interventions can significantly improve anticoagulation management of warfarin [23]. In addition, the IMPACT-AF (Integrated Management Program Advancing Community Treatment of Atrial Fibrillation) study, which investigated the use of oral anticoagulation in patients with AF, found that it was improved by a multifaceted, multilevel (including at medication initiation as well as at follow-up) educational program implemented by physicians [24]. The Smart AF app has educational contents to help patients quickly solve their questions about AF. Thus, the smart AF app can enhance patient education and medication reminders at any time and place. In addition, we believe that we have succeeded in encouraging long-term patient engagement in self-care by avoiding complex operations and time-consuming input as much as possible.

Retention Rate of the Smart App

The decrease in the retention rate of the app over time is an important issue to be considered. The mean retention rates at 1, 3, and 6 months were 0.48 (SD 3), 0.36 (SD 3), and 0.27 (SD 1), respectively. A decline in app use over time is also a concern that has been highlighted in previous reports [25].

Although the app offers the advantage of completely remote recruitment and enrollment, lack of human communication may mean less motivation for the participants to continue compared with studies conducted face-to-face.

The Smart AF app was designed to pop up notifications (reminder) every morning and evening, even when the app had not been activated or opened. Such pop-up feature may have succeeded in improving medication adherence regardless of whether the app was activated or not. However, long-term use is essential for the app to affect users. Clearly, additional efforts to improve retention rates are necessary. The characteristics associated with the retention time of apps have not been well studied. In a study of patients with asthma, being female and older was related to longer retention [26]. In our study, 95 patients enrolled in the study had an HADS-T score of 20 or higher (major anxiety and depression), and most of them had anxiety and depression. As shown in Table 2, older patients, patients with a regular diet, and patients with anxiety had higher retention rates. Conversely, younger patients, those with irregular diets, and those without anxiety had lower retention rates. To deliver mHealth effectively, it is important to identify patient domain factors, such as psychological factors, dietary regularity, and age, that mark suitable candidates for the app.

In the future, to further increase continued app usage rates, we believe that we must work to incorporate an interactive design; more specifically, design elements that respond immediately to patient operations and behaviors.

Limitations

There were limitations to our study and findings. First, this was a single-center study, and the patients were not prospectively randomized into intervention (the app group) and usual care groups. Second, while important in this type of study, the reliance on measurement of self-reported results is a challenge often encountered. The data obtained from these indices may be supported by more objective observations. For example, we used a self-reported medication adherence tool to examine medication adherence. Older adults taking multiple medications may not be able to accurately report their medication use status due to poor memory or confusion. Therefore, it may be useful to complement self-reported measures using more objective measures of medication adherence (eg, medical record review and pharmacy documentation). Third, other new interventions, strategies, and technologies designed to enhance long-term adherence to DOACs need to be developed and investigated because patients with AF are a large and diverse patient population, and not all will have access to newer mHealth tools. Nonadherence is often caused by a multitude of factors, indicating the necessity of providing patients with tailored and more personalized tools. Lastly, this app only showed an improvement in adherence for 6 months; therefore, future studies are needed for long-term improvement.

Conclusion

The Smart AF app improved medication adherence among elderly patients with AF. In the realm of medication management, an approach using an mHealth technology that emphasizes education, automatic reminder, and patient engagement may be helpful. The challenge that emerged, however, was the decline in the rate of persistent use of the app over time; therefore, continuous doctor-patient interaction via the app will be necessary in the future.

Conflicts of Interest

KS is an executive director of the General Incorporated Association of Health Tech Innovation Center, Osaka, Japan.

References

1. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 2014 Feb 25;129(8):837-847 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.113.005119] [Medline: 24345399]

XSL•FO
RenderX

2. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of Current and Future Incidence and Prevalence of Atrial Fibrillation in the U.S. Adult Population. The American Journal of Cardiology 2013 Oct;112(8):1142-1147. [doi: 10.1016/j.amjcard.2013.05.063]
3. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, et al. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2019 Jul 09;74(1):104-132. [doi: 10.1016/j.jacc.2019.01.011] [Medline: 30703431]
4. Jackevicius CA, Tsadok MA, Essebag V, Atzema C, Eisenberg MJ, Tu JV, et al. Early non-persistence with dabigatran and rivaroxaban in patients with atrial fibrillation. Heart 2017 Sep;103(17):1331-1338. [doi: 10.1136/heartjnl-2016-310672] [Medline: 28286333]
5. Deshpande CG, Kogut S, Willey C. Real-World Health Care Costs Based on Medication Adherence and Risk of Stroke and Bleeding in Patients Treated with Novel Anticoagulant Therapy. J Manag Care Spec Pharm 2018 May;24(5):430-439 [FREE Full text] [doi: 10.18553/jmcp.2018.24.5.430] [Medline: 29694285]
6. Ozaki AF, Choi AS, Le QT, Ko DT, Han JK, Park SS, et al. Real-World Adherence and Persistence to Direct Oral Anticoagulants in Patients With Atrial Fibrillation: A Systematic Review and Meta-Analysis. Circ Cardiovasc Qual Outcomes 2020 Mar;13(3):e005969. [doi: 10.1161/CIRCOUTCOMES.119.005969] [Medline: 32148102]
7. Montalescot G, Brotons C, Cosyns B, Crijns HJ, D'Angelo A, Drouet L, AEGEAN Study Investigators. Educational Impact on Apixaban Adherence in Atrial Fibrillation (the AEGEAN STUDY): A Randomized Clinical Trial. Am J Cardiovasc Drugs 2020 Feb;20(1):61-71 [FREE Full text] [doi: 10.1007/s40256-019-00356-2] [Medline: 31243691]
8. Shore S, Ho PM, Lambert-Kerzner A, Glorioso TJ, Carey EP, Cunningham F, et al. Site-level variation in and practices associated with dabigatran adherence. JAMA 2015 Apr 14;313(14):1443-1450. [doi: 10.1001/jama.2015.2761] [Medline: 25871670]
9. Desteghe L, Vijgen J, Koopman P, Dilling-Boer D, Schurmans J, Dendale P, et al. Telemonitoring-based feedback improves adherence to non-vitamin K antagonist oral anticoagulants intake in patients with atrial fibrillation. Eur Heart J 2018 Apr 21;39(16):1394-1403. [doi: 10.1093/eurheartj/ehx762] [Medline: 29300888]
10. Márquez-Contreras E, Martell-Claros N, Márquez-Rivero S, Hermida-Campa E, Gracia-Diez C, Sanchez-López E, ComplianceInertia Working Group, Spanish Society of Hypertension (SEH-LELHA). Strategies for improving dabigatran adherence for stroke prevention in patients with non-valvular atrial fibrillation: education and drug intake reminders (FACILITA study). Curr Med Res Opin 2018 Jul;34(7):1301-1308. [doi: 10.1080/03007995.2018.1435519] [Medline: 29384410]
11. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014;11:CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
12. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]
13. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
14. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Saleem DH, et al. Medication Adherence Apps: Review and Content Analysis. JMIR Mhealth Uhealth 2018 Mar 16;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
15. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile Apps for Medication Management: Review and Analysis. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e13608 [FREE Full text] [doi: 10.2196/13608] [Medline: 31512580]
16. Desteghe L, Kluts K, Vijgen J, Koopman P, Dilling-Boer D, Schurmans J, et al. The Health Buddies App as a Novel Tool to Improve Adherence and Knowledge in Atrial Fibrillation Patients: A Pilot Study. JMIR Mhealth Uhealth 2017 Jul 19;5(7):e98 [FREE Full text] [doi: 10.2196/mhealth.7420] [Medline: 28724512]
17. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GYH. Mobile Health Technology for Atrial Fibrillation Management Integrating Decision Support, Education, and Patient Involvement: mAF App Trial. Am J Med 2017 Dec;130(12):1388-1396.e6 [FREE Full text] [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]
18. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10(5):348-354 [FREE Full text] [Medline: 18453793]
19. Berlowitz DR, Foy CG, Kazis LE, Bolin LP, Conroy MB, Fitzpatrick P, SPRINT Research Group. Effect of Intensive Blood-Pressure Treatment on Patient-Reported Outcomes. N Engl J Med 2017 Aug 24;377(8):733-744 [FREE Full text] [doi: 10.1056/NEJMoa1611179] [Medline: 28834483]
20. Bress AP, Bellows BK, King JB, Hess R, Beddhu S, Zhang Z, SPRINT Research Group. Cost-Effectiveness of Intensive versus Standard Blood-Pressure Control. N Engl J Med 2017 Aug 24;377(8):745-755 [FREE Full text] [doi: 10.1056/NEJMsa1616035] [Medline: 28834469]
21. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, ESC Scientific Document Group. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J 2020 Aug 29:373-498. [doi: 10.1093/eurheartj/ehaa612] [Medline: 32860505]

XSL•FO
RenderX

22. Clarkesmith DE, Lip GYH, Lane DA. Patients' experiences of atrial fibrillation and non-vitamin K antagonist oral anticoagulants (NOACs), and their educational needs: A qualitative study. Thromb Res 2017 May;153:19-27. [doi: 10.1016/j.thromres.2017.03.002] [Medline: 28314139]
23. Clarkesmith DE, Pattison HM, Lip GYH, Lane DA. Educational intervention improves anticoagulation control in atrial fibrillation patients: the TREAT randomised trial. PLoS One 2013;8(9):e74037 [FREE Full text] [doi: 10.1371/journal.pone.0074037] [Medline: 24040156]
24. Vinereanu D, Lopes RD, Bahit MC, Xavier D, Jiang J, Al-Khalidi HR, IMPACT-AF investigators. A multifaceted intervention to improve treatment with oral anticoagulants in atrial fibrillation (IMPACT-AF): an international, cluster-randomised trial. Lancet 2017 Oct 14;390(10104):1737-1746. [doi: 10.1016/S0140-6736(17)32165-7] [Medline: 28859942]
25. Yamaguchi S, Waki K, Nannya Y, Nangaku M, Kadowaki T, Ohe K. Usage Patterns of GlucoNote, a Self-Management Smartphone App, Based on ResearchKit for Patients With Type 2 Diabetes and Prediabetes. JMIR Mhealth Uhealth 2019 Apr 24;7(4):e13204 [FREE Full text] [doi: 10.2196/13204] [Medline: 31017586]
26. Chan YFY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Apr;35(4):354-362 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]

Abbreviations

AEGEAN: Assessment of an Education and Guidance Programme for Eliquis Adherence in Non-Valvular Atrial Fibrillation
AF: atrial fibrillation
CHADS2: congestive heart failure, hypertension, age, diabetes, previous stroke, or transient ischemic attack
DOAC: direct oral anticoagulant
FACILITA: strategies for improving dabigatran adherence for stroke prevention in patients with non-valvular atrial fibrillation
HADS: Hospital Anxiety and Depression Scale
IMPACT-AF: Integrated Management Program Advancing Community Treatment of Atrial Fibrillation
JRISTA: the BMS/Pfizer Japan Thrombosis Investigator Initiated Research Program
mHealth: mobile health
MMAS-8: 8-item Morisky Medication Adherence Scale
OAC: oral anticoagulation

Edited by L Buis; submitted 29.05.21; peer-reviewed by R Ng; comments to author 18.06.21; revised version received 25.11.21; accepted 07.12.21; published 07.01.22.

Please cite as:
Senoo K, Miki T, Ohkura T, Iwakoshi H, Nishimura T, Shiraishi H, Teramukai S, Matoba S
A Smartphone App to Improve Oral Anticoagulation Adherence in Patients With Atrial Fibrillation: Prospective Observational Study
JMIR Mhealth Uhealth 2022;10(1):e30807
URL: https://mhealth.jmir.org/2022/1/e30807
doi:10.2196/30807
PMID:34894626

©Keitaro Senoo, Tomonori Miki, Takashi Ohkura, Hibiki Iwakoshi, Tetsuro Nishimura, Hirokazu Shiraishi, Satoshi Teramukai, Satoaki Matoba. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 07.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adherence to Growth Hormone Treatment Using a Connected Device in Latin America: Real-World Exploratory Descriptive Analysis Study

Aria Assefi[1], MD; Paula van Dommelen[2], PhD; Lilian Arnaud[3], MSc; Carlos Otero[4], MD; Luis Fernandez-Luque[5], PhD; Ekaterina Koledova[6], MD, PhD; Luis Eduardo Calliari[7], MD

[1]Fertility and Endocrinology, Merck SA (an affiliate of Merck KGaA, Darmstadt, Germany), Buenos Aires, Argentina
[2]Department of Child Health, The Netherlands Organization for Applied Scientific Research TNO, Leiden, Netherlands
[3]Global Healthcare Operations, Connected Health & Devices, Ares Trading SA (an affiliate of Merck KGaA, Darmstadt, Germany), Eysins, Switzerland
[4]Departamento de Informática en Salud, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina
[5]Adhera Health Inc, Palo Alto, CA, United States
[6]Global Medical Affairs Cardiometabolic & Endocrinology, the healthcare business of Merck KGaA, Darmstadt, Germany
[7]Pediatric Endocrinology Unit, Pediatric Department, Santa Casa de São Paulo School of Medical Sciences, São Paulo, Brazil

Corresponding Author:
Aria Assefi, MD
Fertility and Endocrinology
Merck SA (an affiliate of Merck KGaA, Darmstadt, Germany)
Edificio Panamericana Plaza
Tronador
Buenos Aires, 4890
Argentina
Phone: 54 9 11 2275146
Email: aria.assefi@merckgroup.com

Abstract

Background: Recombinant human growth hormone (rhGH) therapy is an effective treatment for children with growth disorders. However, poor outcomes are often associated with suboptimal adherence to treatment.

Objective: The easypod connected injection device records and transmits injection settings and dose data from patients receiving rhGH. In this study, we evaluated adherence to rhGH treatment, and associated growth outcomes, in Latin American patients.

Methods: Adherence and growth data from patients aged 2-18 years from 12 Latin American countries were analyzed. Adherence data were available for 6207 patients with 2,449,879 injections, and growth data were available for 497 patients with 2232 measurements. Adherence was categorized, based on milligrams of rhGH injected versus milligrams of rhGH prescribed, as high (≥85%), intermediate (>56%-<85%), or low (≤56%). Transmission frequency was categorized as high (≥1 per 3 months) or low (<1 per 3 months). Chi-square tests were applied to study the effect of pubertal status at treatment start and sex on high adherence, and to test differences in frequency transmission between the three adherence levels. Multilevel linear regression techniques were applied to study the effect of adherence on observed change in height standard deviation score (ΔHSDS).

Results: Overall, 68% (4213/6207), 25% (n=1574), and 7% (n=420) of patients had high, intermediate, and low adherence, respectively. Pubertal status at treatment start and sex did not have a significant effect on high adherence. Significant differences were found in the proportion of patients with high transmission frequency between high (2018/3404, 59%), intermediate (608/1331, 46%), and low (123/351, 35%) adherence groups (P<.001). Adherence level had a significant effect on ΔHSDS (P=.006). Mean catch-up growth between 0-24 months was +0.65 SD overall (+0.52 SD in patients with low/intermediate monthly adherence and +0.69 SD in patients with high monthly adherence). This difference translated into 1.1 cm greater catch-up growth with high adherence.

Conclusions: The data extracted from the easypod Connect ecosystem showed high adherence to rhGH treatment in Latin American patients, with positive growth outcomes, indicating the importance of connected device solutions for rhGH treatment in patients with growth disorders.

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2022;10(1):e32626) doi:10.2196/32626

KEYWORDS

adherence; connected device; growth; growth disorders; growth hormone; electronic injection device; real-world data; disorder; therapy; treatment; children; outcome

Introduction

Adherence to long-term pharmacological treatments, such as growth hormone (GH) therapy for growth disorders, is an area with great potential for improvement [1,2]. Poor long-term adherence to GH treatment is known to affect final adult height and additional clinical outcomes in children with growth disorders [3]. Moreover, enthusiasm and motivation to adhere to treatment may decrease over time because the long-term benefits of GH treatment are not immediately obvious to children, and administering daily subcutaneous injections places a significant burden on them and their parents/caregivers [3-5]. Indeed, adherence to GH treatment has been shown to be statistically significantly higher in treatment-naive children compared with those experienced in their treatment [4].

Adherence to GH treatment in the real-world setting has always been difficult to monitor, given the use of unreliable proxy methods such as patient testimony or records of prescriptions filled/vials counted [4,5]. Furthermore, detection of poor adherence to GH treatment can be problematic because patients/caregivers may be reluctant to admit to (or do not remember) missed doses and may overestimate their adherence to treatment during discussions with health care providers (HCPs) [4]. Devices that offer a dose-setting memory may therefore be beneficial in improving adherence [6]. In addition, with prevalence estimates of nonadherence ranging from 5% to 82% [5], it is difficult to compare adherence rates among studies due to the variability in methods used to evaluate and define adherence [5,7]. Lastly, studies assessing adherence are sometimes constrained by low patient numbers and there is often only one participating center, which limits the extrapolation of results to different settings [8].

Automatic transmission of injection data provides a more accurate insight into real-world adherence patterns and enables HCPs to potentially eliminate poor adherence as a reason for a suboptimal response to GH treatment [4,9,10]. The use of a connected injection device to deliver GH treatment limits the risk of misreporting or faulty recall of adherence, and allows HCPs to accurately monitor their patients' real-world adherence behavior over time [9]. Patient confidentiality is maintained because only the treating HCPs and patient support programs (PSPs) can access patients' complete data from a secure cloud-based database, and only deidentified (pseudonymized) data are used to generate aggregated and anonymized results for research purposes. A connected injection device thereby enables HCPs to access the transmitted data and gain insights into both individual and overall patterns of adherence to GH treatment [9,11].

As different health care systems and variations in clinical practice around the world may affect adherence locally, the deployment of a connected injection device for the treatment of growth disorders across different countries allows the study of behavioral adherence patterns across populations and longitudinally for thousands of patients. This substantial compendium of patient-generated data has also been applied in the context of understanding patterns in diabetes thanks to glucose monitoring technologies [12,13]. In terms of monitoring adherence in large cohorts, oral medication use has been monitored using "smart pillboxes" [14] and "smart pill bottles" [15,16].

Global analysis of real-world data obtained from connected injection devices for GH treatment has shown that children with high adherence were most likely to regularly transmit data, and that prepubertal children showed higher adherence than older children and adolescents [17]. This analysis showed the potential of developing a global adherence decision support system (ADSS) by analyzing trends in real-world adherence data, but did not include insights from different health care systems.

The Latin American region is one of the fastest growing regions in terms of the adoption of digital health [18,19]. This has facilitated the implementation of a web-based platform connected with injection devices for GH treatment across several countries in Latin America, a region in which digital health studies have previously been lacking. The objective of this study, therefore, was to evaluate real-world adherence to recombinant human GH (rhGH) therapy administered via a connected injection device in one specific region (Latin America) to provide an update to an earlier, smaller Latin American analysis, previously published only in abstract form [11]. Additionally, we studied catch-up growth and its association with adherence in a subgroup of patients.

Methods

Patient Population

In this analysis, we included children with growth disorders from 12 Latin American countries (Argentina, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, El Salvador, Guatemala, Mexico, Nicaragua, Panama, and Peru) and assessed the effects of pubertal status at treatment start, sex, and engagement with treatment on their adherence. An analysis of longitudinal records for a total of 13,553 children in the global database was recently published [17], but here we focus only on data from patients in these Latin American countries. The total data set available in the global database comprises data from the 5-year global Easypod Connect Observational Study (ECOS) [9] and from all patients worldwide who have received treatment with rhGH (somatropin [Saizen]; the healthcare business of Merck Healthcare KGaA, Darmstadt, Germany) administered via the easypod connected injection device.

XSL•FO
RenderX

A Digital Ecosystem for Supporting and Monitoring Adherence to rhGH Therapy

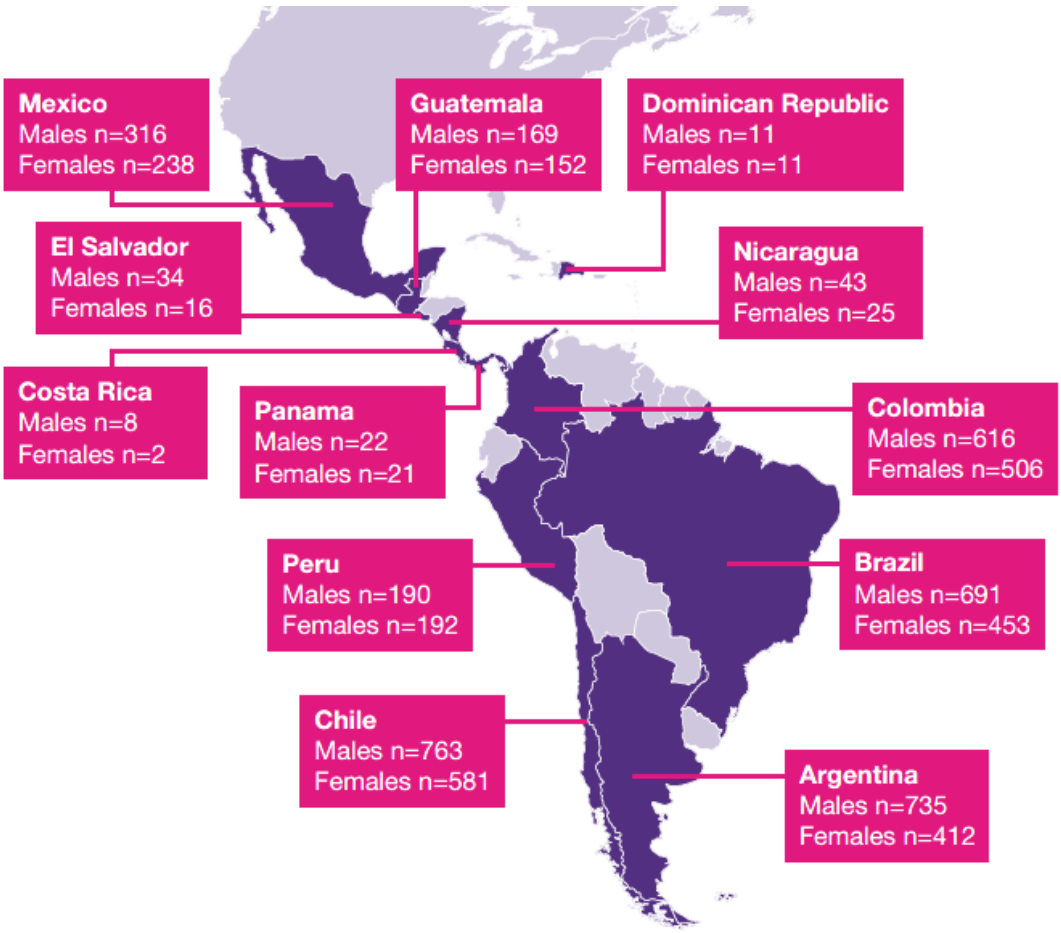
The easypod electromechanical injection device, in combination with the easypod Connect web-based platform, as part of an ADSS [20], electronically records accurate, objective details of the date, time, and dose of injections for patients receiving rhGH for the treatment of growth disorders [21]. All of these data are recorded by the device and stored internally for up to 3 years. The patient can then transmit these data to the easypod Connect platform and a secure internet cloud-based database [9].

Study Design and Inclusion Criteria

This was an exploratory descriptive analysis study during which 4 years of adherence data were analyzed from 6207 pediatric patients with 2,449,879 prescribed injections of rhGH delivered via easypod to treat growth disorders, and who were transmitting data to the easypod Connect system between January 2007 and December 2020. These patients resided in 12 Latin American countries (Figure 1). To avoid inclusion of test doses or training injections, only data after the 10th injection registered for each individual were analyzed. Data were downloaded from the easypod Connect platform in January 2021, but the period of recorded data varied according to each individual patient's treatment duration. Transmission data from January 2016 to December 2020 were used. We selected patients who were aged 2-18 years at treatment start.

Eligible patients from each of the participating countries had been enrolled in the database and attended at least one visit, according to local routine clinical practice. Diagnoses and decisions on treatment were made at the discretion of the physician responsible for each patient, following standard endocrinologic practice for each of the participating countries. Prior to enrollment, patients/caregivers reviewed and voluntarily signed an informed consent form materializing their agreement for data collection, storage, and use of their child's pseudonymized data to create aggregated statistical and general adherence reports.

Figure 1. Participating Latin American countries.



For height, we selected patients with at least two measurements. Adherence (calculated as milligrams of rhGH injected versus milligrams of rhGH prescribed) was categorized as high (≥85%), intermediate (>56%-<85%), or low (≤56%) for patients on either 6 or 7 injections per week, the two possible regimens for treatment with rhGH. The dosage and frequency of rhGH therapy as per easypod settings were defined by HCPs and data transmissions were initiated by the child, parent/caregiver, or HCP. Adherence was assessed overall and monthly, and explored by puberty status at treatment start (nominal cutoffs at 10 years for girls, and 12 years for boys), sex, and transmission frequency (defined as the total number of transmissions divided by the duration of treatment, and categorized as high [≥1 per 3 months] versus low [<1 per 3 months]) in the selected patients with available adherence data for ≥3 months between January 2016 and December 2020.

XSL•FO
RenderX

Transmission frequency was calculated in each adherence category as a proxy measure of the patient's engagement in disease management using easypod. No imputation was made for missing data or withdrawal from the study.

Patient Data and Calculations

Height data were available for 497 patients with 2232 measurements; this included 355 patients from Argentina, 64 from Brazil, 70 from Guatemala, and 8 from Mexico. Height standard deviation scores (HSDS) were calculated using the World Health Organization references [22,23]. Linear interpolation between height measurements was applied to calculate monthly catch-up growth (ΔHSDS) overall, and by low/intermediate versus high adherence, and within the subgroup of patients (n=40) aged <8 years with HSDS of <–2, for which we assume an optimal catch-up growth. The cutoff of 8 years was chosen so the definition of short stature, age, and treatment duration ensured this would be before the start of puberty in both genders. Cubic smoothing splines were fitted to obtain the curves for catch-up growth between 0-24 months.

Statistical Analysis

Descriptive statistics were used to describe differences over time in adherence (low, intermediate, or high), puberty status (prepubertal or pubertal), and sex. Chi-square tests were applied to test differences in high adherence between girls and boys, and between prepubertal and pubertal girls and boys. In addition, a Chi-square test was applied to test differences in high frequency transmission between the high, intermediate, and low adherence groups. Multilevel linear regression techniques were applied to study the effect of adherence level on ΔHSDS between all observed growth measurements, adjusted for the time intervals between them.

Ethical Considerations

Treatment with rhGH via easypod was conducted according to local practice. This real-world, retrospective analysis of the data set was performed in accordance with the informed consent form, signed by caregivers of children and adult patients materializing their agreement for data collection, storage, and use of their pseudonymized data to create aggregated statistical and general adherence reports.

Data Availability Statement

Any requests for data by qualified scientific and medical researchers for legitimate research purposes will be subject to the healthcare business of Merck KGaA's Data Sharing Policy. All requests should be submitted in writing to the healthcare business of Merck KGaA's data sharing portal [24]. When the healthcare business of Merck KGaA has a coresearch, codevelopment, or comarketing or copromotion agreement, or when the product has been outlicensed, the responsibility for disclosure might be dependent on the agreement between parties. Under these circumstances, the healthcare business of Merck KGaA will endeavor to gain agreement to share data in response to requests.

Results

Patient Population and Demographics

Complete data were available for 6207 patients, where "overall" is defined here (and throughout) as the total number of patients who received rhGH, started treatment at age 2-18 years, and for whom data were available. Transmission data were available for 5086 patients. Patient demographics according to adherence rates are presented for this data set in Table 1. The number of patients decreased from 6207 patients in the first month (100%), to 3594 patients (58%) at month 12, to 1707 patients at month 24, and to <600 patients (<10%) after month 36 (Figure 2; black line). Growth data were available for 497 patients overall, and decreased to 330 patients at months 13-24, 150 patients at months 25-36, and 37 patients at months 37-48.

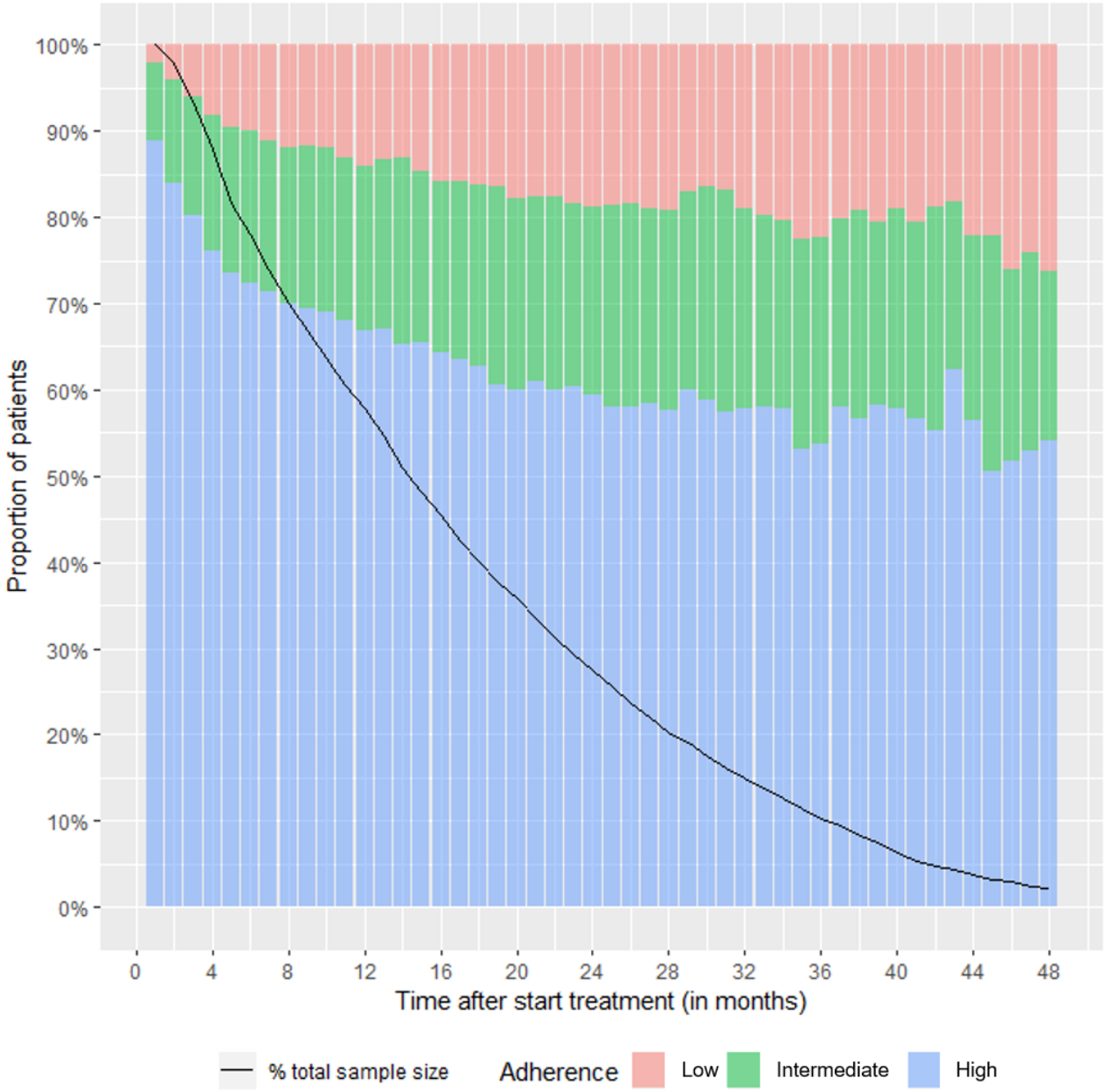
Table 1. Overall patient demographics according to adherence rates.

Characteristic	Adherence[a]			Total (N=6207)
	High (n=4213)	Intermediate (n=1574)	Low (n=420)	
Mean age (SD) of boys at start, years	10.6 (3.4)	10.9 (3.3)	10.6 (3.4)	10.7 (3.3)
Boys aged <12 years at start, n (%)	1424 (68)	526 (25)	157 (7)	2107 (34)
Boys aged ≥12 years at start, n (%)	985 (66)	394 (26)	112 (8)	1491 (24)
Mean age (SD) of girls at start, years	9.7 (2.8)	10.0 (2.8)	9.8 (3.2)	9.8 (2.8)
Girls aged <10 years at start, n (%)	839 (69)	308 (25)	72 (6)	1219 (20)
Girls aged ≥10 years at start, n (%)	965 (69)	346 (25)	79 (6)	1390 (22)

[a]Adherence was categorized as high (≥85%), intermediate (>56%-<85%), or low (≤56%).

XSL•FO
RenderX

Figure 2. Proportion of patients who were adherent at each time point. Adherence was recorded for the cross-section of children or their caregivers transmitting data at each time point; no imputation was made for missing data or withdrawal from the study.



Adherence Overall and Over Time

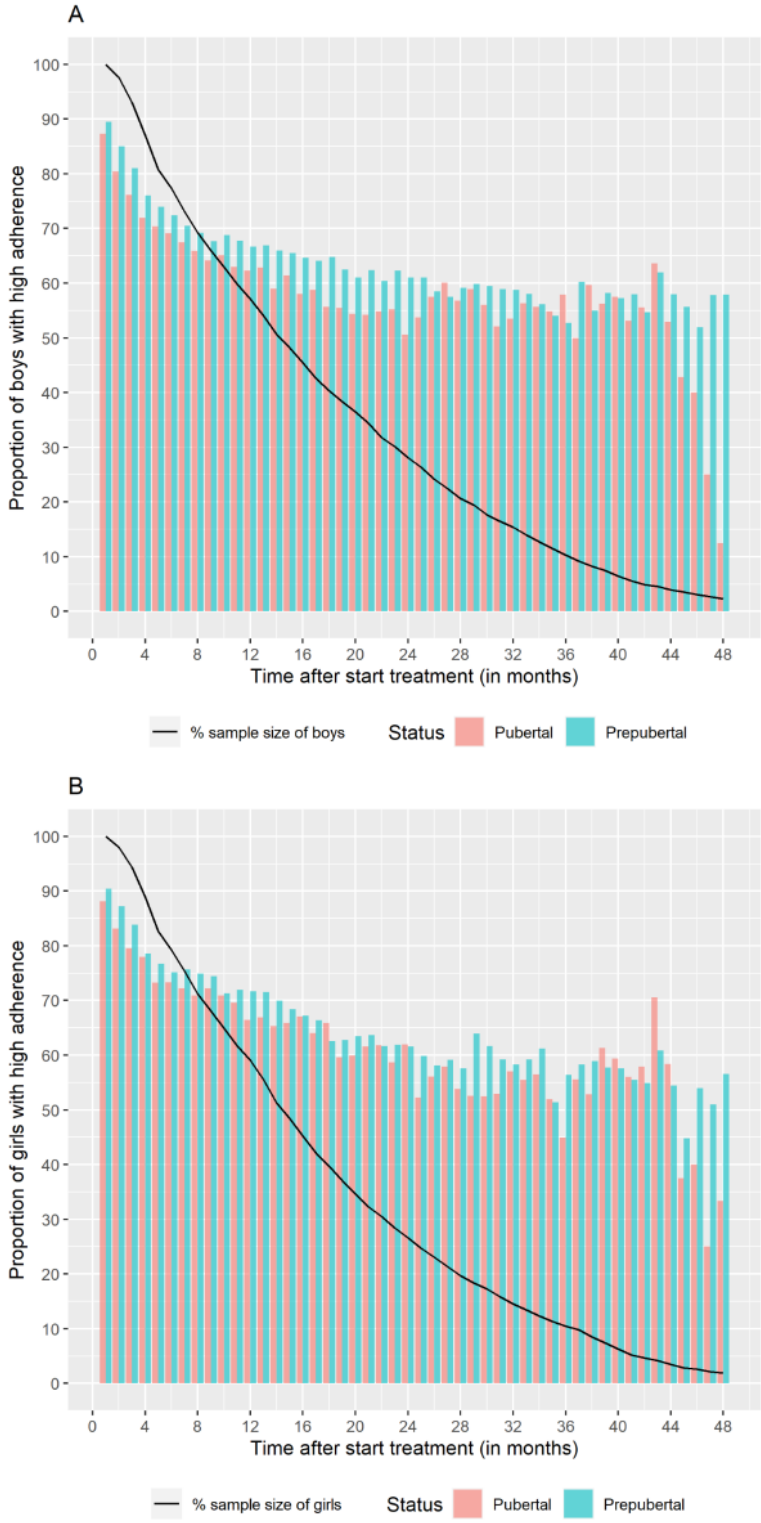
Overall, 68% (4213/6207) of patients were in the high adherence category, 25% (n=1574) were in the intermediate adherence category, and 7% (n=420) were in the low adherence category. Furthermore, at each time point, there was a higher proportion of patients in the high adherence category than in the intermediate and low categories combined (Figure 2). High adherence decreased from 89% (5514/6207) to 59% (1013/1707) between 1-24 months, and ranged from 50%-62% between 25-48 months. However, despite there being a decrease in the proportion of patients in the high adherence category over time, 67% (2399/3594) and 59% (1013/1707) of patients were still in the high adherence category at months 12 and 24, respectively.

Effect of Age, Sex, and Engagement With the Easypod Device on Adherence

There were no significant differences in high adherence between boys (2409/3598, 67%) and girls (1804/2609, 69%) or between prepubertal and pubertal patients (1424/2107, 68% versus 985/1491, 66% in boys, both 69% [839/1219; 965/1390] in girls, respectively). Figures 3A and 3B show adherence at each time point stratified by nominal pubertal age and sex. Figure 3A shows a larger proportion of prepubertal boys with high adherence between 1-25 months compared with pubertal boys. For the majority of these months (17/25), the proportion of high adherence was significantly (P<.05) higher in prepubertal boys compared with pubertal boys.

XSL•FO
RenderX

Figure 3. Adherence at each time point stratified by nominal pubertal age and sex (A: boys, B: girls). Adherence was recorded for the cross-section of children or their caregivers transmitting data at each time point; no imputation was made for missing data or withdrawal from the study.



In addition, there were significant differences in the proportion of patients with high transmission frequency between the adherence groups; 59% (2018/3404), 46% (608/1331), and 35% (123/351) in the high, intermediate, and low adherence groups, respectively (P<.001).

Between-Country Variation

The proportion of children in the high adherence category varied by country, ranging from 45% (10/22) in the Dominican Republic to 82% (943/1144) in Brazil (Table 2).

XSL•FO
RenderX

Table 2. Adherence rates by age, sex, and country.

Country and adherence[a]	Boys aged <12 years, n	Boys aged ≥12 years, n	Girls aged <10 years, n	Girls aged ≥10 years, n	Total number of patients with ≥1 injection	Adherence ≥85%, n (%)
Argentina					1147	759 (66)
High	322	163	151	123		
Intermediate	113	72	66	41		
Low	46	19	18	13		
Brazil					1144	943 (82)
High	347	225	169	202		
Intermediate	68	39	35	35		
Low	9	3	6	6		
Chile					1344	949 (71)
High	291	242	176	240		
Intermediate	96	84	44	95		
Low	17	33	12	14		
Colombia					1122	625 (56)
High	170	152	149	154		
Intermediate	125	108	81	86		
Low	39	22	18	18		
Costa Rica					10	8 (80)
High	4	3	0	1		
Intermediate	0	1	1	0		
Low	0	0	0	0		
Dominican Republic					22	10 (45)
High	1	3	3	3		
Intermediate	2	0	0	4		
Low	2	3	0	1		
El Salvador					50	27 (54)
High	7	12	5	3		
Intermediate	8	5	2	3		
Low	1	1	2	1		
Guatemala					321	221 (69)
High	71	38	43	69		
Intermediate	26	14	18	16		
Low	8	12	0	6		
Mexico					554	355 (64)
High	113	83	72	87		
Intermediate	48	40	32	33		
Low	23	9	6	8		
Nicaragua					68	33 (49)
High	13	8	5	7		
Intermediate	8	7	5	4		
Low	6	1	2	2		
Panama					43	23 (53)

XSL•FO
RenderX

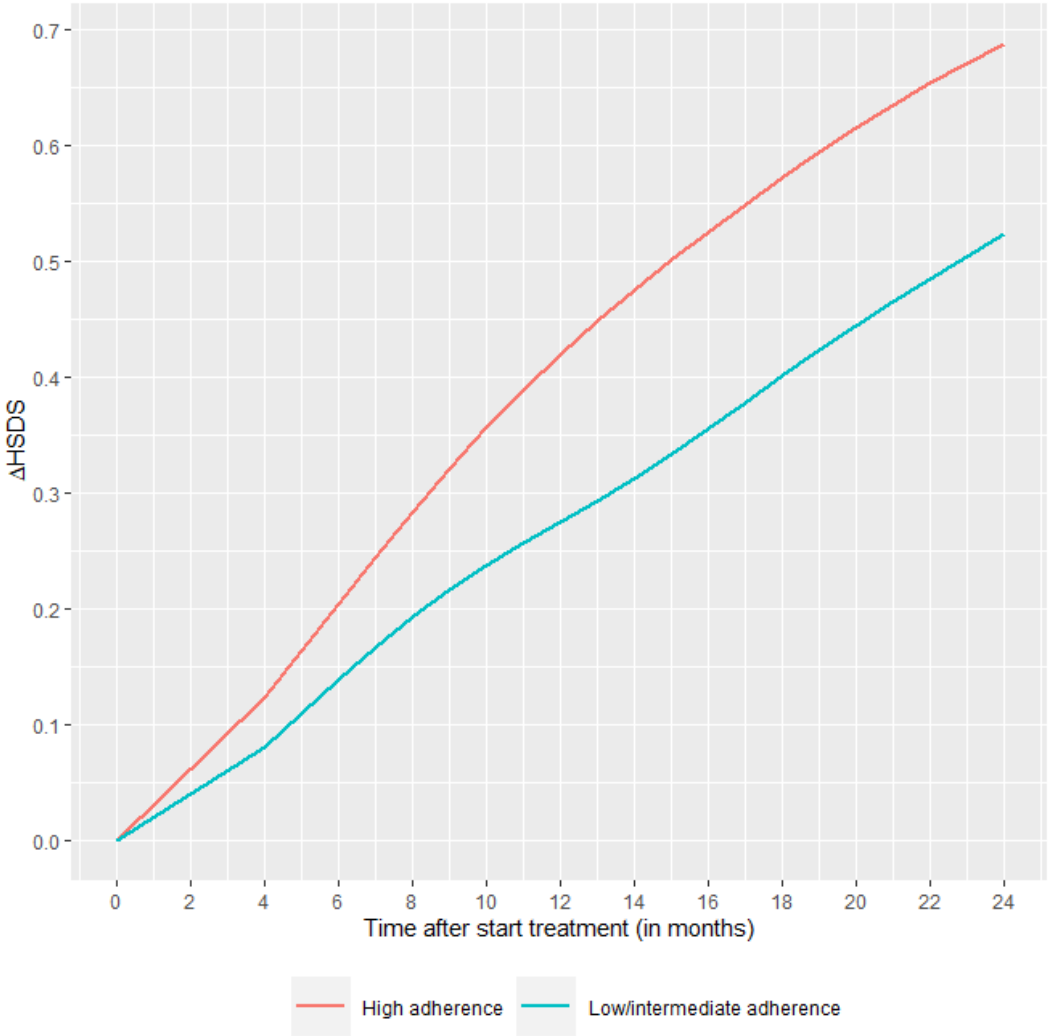
Country and adherence[a]	Boys aged <12 years, n	Boys aged ≥12 years, n	Girls aged <10 years, n	Girls aged ≥10 years, n	Total number of patients with ≥1 injection	Adherence ≥85%, n (%)
High	6	5	4	8		
Intermediate	5	3	3	6		
Low	2	1	0	0		
Peru					382	260 (68)
High	79	51	62	68		
Intermediate	27	21	21	23		
Low	4	8	8	10		

[a]Adherence was categorized as high (≥85%), intermediate (>56%-<85%), or low (≤56%).

Catch-up Growth

Figure 4 shows the mean catch-up growth (ΔHSDS) between 0-24 months, stratified by high versus intermediate/low adherence. Adherence level (low/intermediate versus high) had a significant effect on ΔHSDS (*P*=.006). Mean catch-up growth between 0-12 months was +0.39 SD overall, with +0.27 SD in patients with low/intermediate monthly adherence and +0.42 SD in patients with high monthly adherence. Mean catch-up growth between 0-24 months was +0.65 SD overall, with +0.52 SD in patients with low/intermediate monthly adherence and +0.69 SD in patients with high monthly adherence. Mean catch-up growth within the subgroup of patients (n=40) aged <8 years and HSDS <–2 at start was +1.03 SD. It was not possible to stratify this subgroup of patients by low and high adherence due to the small sample size.

Figure 4. Catch-up growth between 0-24 months stratified by low/intermediate adherence (<85%) and high adherence (≥85%). ΔHSDS: change in height standard deviation score.



XSL•FO
RenderX

Discussion

Principal Results

The data extracted from the easypod Connect ecosystem showed that children receiving rhGH therapy demonstrated high adherence over 4 years in a real-world setting in 12 Latin American countries.

Use of easypod Connect, which showed high adherence in patients with high transmission rates, could be regarded as a proxy of usability and utility of the system by patients and HCPs. In such large-scale deployments of digital health systems, usability cannot be easily measured using questionnaires as this would pose a major user/participant burden. Instead, transmissions and data usage can be used toward identifying patterns of usage of the system and adherence to treatment. Indeed, in this regard, a framework has been proposed to guide future implementation and research on the use of digital health tools to support patients with growth disorders who require GH therapy [25]. These insights can be further investigated in smaller-scale studies (eg, exploring qualitative issues with interviews).

We observed that patients who transmitted their data tended to have higher adherence; therefore, a potential hypothesis to further explore is the impact of patients knowing that their adherence data is being observed and is useful to their HCP. Studies in other therapeutic areas have found that factors related to the Theory of Planned Behavior can predict ~50% variance in adherence [26]. Future studies can look into the behavioral impact of monitoring adherence and the role this plays in the interaction of HCPs with the data during clinical visits (eg, clinicians reviewing adherence data during the patient visit, how the system is introduced to the patient).

Comparison With Other Studies

The selected data from this regional study are broadly consistent with recently published data from the global easypod Connect database, which demonstrated that, of 13,553 children, 71% were in the high adherence category [17], and with a previous exploratory analysis (68% in the high adherence category) [27]. Children with high adherence were most likely to regularly transmit data, reflecting engagement with the easypod device [17]. Furthermore, the ECOS has shown that the majority (62%) of patients maintained an observed adherence rate of ≥80% to rhGH therapy over 3 years of easypod use [9]. Similarly, a preliminary analysis evaluated real-world adherence to rhGH therapy administered via easypod over 1, 3, 6, and 12 months in Latin American patients [11]. Analysis of data extracted in February 2018 from 2727 patients transmitting their injection data to easypod Connect showed that the majority (67%) were still in the high adherence category (defined as ≥85%) at 12 months, with girls (64% versus 63% of boys) and prepubertal patients (69% versus 60% of pubertal patients) being the most adherent [11].

In both the preliminary Latin American study and other analyses, the proportion of children with high adherence declined over time, and factors such as sex and nominal age at puberty appeared to have only a small effect on adherence rates [11,17]. Similarly, in this study, after 12 months, 33% (n=1195 out of N=3594) of Latin American patients were in the low or intermediate adherence categories, while this was 11% (n=693 out of N=6207) in the first month of treatment. Additionally, a slightly higher proportion of girls overall were in the high adherence category compared with boys, as were patients who were prepubertal at treatment start versus pubertal patients; however, overall, these results were not statistically significant. These results demonstrate that adherence is an issue that needs to be addressed continuously by HCPs and included in discussions with patients and their families/caregivers, taking both patient sex and pubertal status into account.

Patient attrition over time may be due to a number of factors. Patients may have been switched to a different rhGH, may have continued taking somatropin rhGH using a pen injector, may have continued taking rhGH using easypod but without performing any further data transmission, or may have stopped treatment completely. Minimizing patient attrition is an important consideration in long-term clinical trials, and determining the predictors of attrition is key to identifying patients at risk of missed visits or dropout; such patients may be excluded from a trial or efforts may need to be made to prevent their subsequent dropout once enrolled [28]. This is especially important in the case of growth disorders since rhGH treatment is required over the long term and requires good adherence to achieve optimal outcomes. Indeed, determining predictors of attrition and using early trial retention strategies (eg, management of reluctant or hard-to-locate study participants) have led to improved attrition rates in studies of asthma and behavioral disorders [28,29]. Thus, such approaches could be considered in future studies investigating adherence to rhGH treatment.

Indeed, previously reported individual cases of patients receiving rhGH have indicated that direct access to adherence monitoring by HCPs followed by intervention can make a difference to a patient's management and motivation [30-32]. The authors' personal experience shows that, in many cases, families are not always aware that their child's adherence is suboptimal and are surprised by this information when informed by the HCP of the data recorded by easypod. This information is key because it can explain a smaller growth catch-up without the need for further investigation into other potential causes. In this regard, devices with a dose setting were the preferred choice among patients and caregivers in a recent study by Tanaka et al [6].

Optimizing adherence might also be achieved through the use of structured and active interventions from HCPs, patient/caregiver support programs [33], and/or digital interventions to help manage adherence over the long-term course of rhGH treatment. Although extensive evidence is available in the adult population [34-36], there are few studies addressing the unique needs of digital adherence support in pediatrics [37]. However, the potential of gamified interventions to promote and improve adherence in pediatric patients has recently been demonstrated [38].

Strengths and Limitations

The strengths of our study include the large data set from a real-world study conducted across 12 countries in one

XSL•FO
RenderX

geographic region with substantially diverse and dynamic health care systems, and the fact that the data are derived from a connected injection device, which offers more reliable data compared to data based on the declarations of patients or their parents/caregivers. Comparable insights into patient adherence from other large-scale patient registries for rhGH treatment are not available due to the lack of alternative electronic devices similar to easypod.

Further work is required to assess whether patient/caregiver engagement, as measured by the rate and frequency of data transmission with easypod, is associated with better long-term adherence and clinical outcomes for patients. It would also be interesting to investigate whether or not there is a correlation between the frequency of dose adjustments and the adherence/transmission rate.

Limitations of the study include patient attrition, summarized above, and the fact that the change in adherence rate over time varied from child to child, perhaps due to changes in individual treatment plans or other actions taken by the HCP or child/family. Furthermore, not all data were available for all patients over the same treatment duration, as would be expected in any observational study. Differences between the 12 participating countries in terms of socioeconomic factors, such as reimbursement and/or out-of-pocket expenses for GH therapy or free access to medication, may also have affected individual or local adherence rates. Similarly, differences in prescribing habits (eg, prescription provided for 1, 2, 3, or 6 months), patients' visits to the clinic where growth response is checked and dose might be adjusted to maximize response to treatment, and refill of prescriptions from country to country may also affect adherence rates. Other potential limitations include the lack of additional information on diagnosis and clinical background, limited data on growth outcomes that do not allow assessment of the full catch-up growth pattern, and lack of patient-reported data such as reasons for discontinuation or interruption of treatment, or predefined actions taken by PSPs [39] in the various countries involved. These data might have been available through linkage to electronic health records (EHRs) where this is permitted, or by allowing patients to have self-reported outcomes entered separately into the system via an app.

Finally, usage of the system (ie, adherence and transmission rates) can be a proxy of usability and utility; however, usage has been to a large extent influenced by the way the system was introduced. All of these areas require more research. Our analysis showed less catch-up growth (–0.17 SD over 24 months) in patients who continue to have low/intermediate adherence compared with patients with high adherence, and this difference increased over time. This difference could be translated into centimeters, resulting in 1.1 cm less catch-up growth between 0-24 months for a patient with an average age for the group (10 years). This is in agreement with the literature, which shows that greater adherence leads to improved growth, with poor adherence adversely affecting growth outcomes [9,10,40]. In the ECOS analyses, statistically significant correlations were observed between adherence and 1-year ΔHSDS ($P<.001$ for patients overall) and height velocity ($P<.001$) [9].

Future Work

In terms of future research opportunities, analysis of the potential differences between countries with preset additional data collection points would be of interest. The value of this data can be enhanced by complementing it with self-reported height data entered via a patient app. Future research may also allow for automatic self-measurement of height using novel augmented reality technology on mobile phones. Finally, integration with EHRs may facilitate clinical workflows, as has been demonstrated in other therapy areas [16].

Conclusions

Analysis of the data extracted from easypod Connect showed high adherence to rhGH treatment in Latin American patients, with positive growth outcomes. Thus, our study indicates the potential value of using a connected injection device to monitor and study adherence at an international level. It shows that through our validated method of recording adherence with easypod, we can address an unmet need in rhGH therapy, enabling HCPs to accurately identify patients for whom interventions to improve adherence would be beneficial to improve their growth and other clinical outcomes, particularly as the proportion of children with high adherence declined over time, which is consistent with previous findings. The data also show that children who were most adherent to treatment were more likely to transmit their injection data results regularly and have larger catch-up growth than those who were less adherent. This association between adherence and transmission of data may indicate that sharing data with HCPs has a positive impact on adherence rates, and further studies to confirm this are needed. High adherence and transmission rates may reflect the positive use of the system and could be regarded as indirect indicators of usability and utility of the system by patients and HCPs.

Acknowledgments

This study was funded by the healthcare business of Merck KGaA, Darmstadt, Germany (CrossRef Funder ID: 10.13039/100009945).

The authors would like to thank the patients and their families and the health care professionals who participated in this study. They would also like to thank Dr Ignacio Bergadá of Hospital de Niños Ricardo Gutiérrez, Buenos Aires, Argentina, for his contribution to the review and interpretation of the data and revision of the manuscript. Medical writing assistance was provided by David Candlish and Amy Evans of inScience Communications, Springer Healthcare Ltd, and funded by the healthcare business of Merck KGaA, Darmstadt, Germany, in accordance with Good Publication Practice (GPP3) guidelines.

XSL•FO
RenderX

Authors' Contributions

AA contributed to review/interpretation of the manuscript. PvD was involved in data analysis, and in data review/interpretation and revision of the manuscript. LA was involved in data collection/analysis and revision of the manuscript. CO contributed to the analysis of the data set and the results of the research project from the perspective of health informatics, evaluation of the evidence for or against the project, and bibliographic review from the perspective of health informatics. LFL was involved in the study design, data interpretation, and revision and review of the manuscript. EK was involved in the concept design, data analysis design, data review/interpretation, and revision of the manuscript. LEC was involved in the design of the study, data review/interpretation, and revision of the manuscript. All authors have approved the final version of the manuscript.

Conflicts of Interest

AA is an employee of Merck SA, Buenos Aires, Argentina (an affiliate of Merck KGaA, Darmstadt, Germany). PvD has a consultancy agreement with the healthcare business of Merck KGaA, Darmstadt, Germany. LA is an employee of Ares Trading SA (an affiliate of Merck KGaA, Darmstadt, Germany). CO does not have any conflicts of interest to declare. LFL is Chief Scientific Officer at Adhera Health Inc, which has a commercial relationship with the healthcare business of Merck KGaA, Darmstadt, Germany. EK is an employee of the healthcare business of Merck KGaA, Darmstadt, Germany, and holds shares in the company. LEC has received honoraria as a lecturer from the healthcare business of Merck KGaA, Darmstadt, Germany, NovoNordisk, and Pfizer.

References

1. Hommel K, Ramsey R, Rich K, Ryan J. Adherence to Pediatric Treatment Regimens. In: Roberts MC, Steele RG, editors. Handbook of Pediatric Psychology, Fifth ed. New York: Guilford Press; 2017:119-133.
2. Adherence to long-term therapies: evidence for action. World Health Organization. 2003. URL: https://www.who.int/chp/knowledge/publications/adherence_report/en/ [accessed 2021-03-01]
3. Cutfield WS, Derraik JGB, Gunn AJ, Reid K, Delany T, Robinson E, et al. Non-compliance with growth hormone treatment in children is common and impairs linear growth. PLoS One 2011 Jan 31;6(1):e16223 [FREE Full text] [doi: 10.1371/journal.pone.0016223] [Medline: 21305004]
4. Bozzola M, Colle M, Halldin-Stenlid M, Larroque S, Zignani M, easypod™ survey study group. Treatment adherence with the easypod™ growth hormone electronic auto-injector and patient acceptance: survey results from 824 children and their parents. BMC Endocr Disord 2011 Feb 04;11:4 [FREE Full text] [doi: 10.1186/1472-6823-11-4] [Medline: 21294891]
5. Fisher BG, Acerini CL. Understanding the growth hormone therapy adherence paradigm: a systematic review. Horm Res Paediatr 2013;79(4):189-196 [FREE Full text] [doi: 10.1159/000350251] [Medline: 23635797]
6. Tanaka T, Sato T, Yuasa Mhwm A, Akiyama T, Tawseef A. Patient preferences for growth hormone treatment in Japanese children. Pediatr Int 2021 Apr 30;63(10):1185-1191. [doi: 10.1111/ped.14760] [Medline: 33930225]
7. Rodríguez Arnao MD, Rodríguez Sánchez A, Díez López I, Ramírez Fernández J, Bermúdez de la Vega JA, Yeste Fernández D, et al. Adherence and long-term outcomes of growth hormone therapy with easypod™ in pediatric subjects: Spanish ECOS study. Endocr Connect 2019 Sep 01;8(9):1240-1249 [FREE Full text] [doi: 10.1530/EC-19-0325] [Medline: 31484160]
8. Maggio MC, Vergara B, Porcelli P, Corsello G. Improvement of treatment adherence with growth hormone by easypod™ device: experience of an Italian centre. Ital J Pediatr 2018 Sep 27;44(1):113 [FREE Full text] [doi: 10.1186/s13052-018-0548-z] [Medline: 30261918]
9. Koledova E, Stoyanov G, Ovbude L, Davies PSW. Adherence and long-term growth outcomes: results from the easypod connect observational study (ECOS) in paediatric patients with growth disorders. Endocr Connect 2018 Aug;7(8):914-923 [FREE Full text] [doi: 10.1530/EC-18-0172] [Medline: 29976785]
10. van Dommelen P, Koledova E, Wit JM. Effect of adherence to growth hormone treatment on 0-2 year catch-up growth in children with growth hormone deficiency. PLoS One 2018;13(10):e0206009 [FREE Full text] [doi: 10.1371/journal.pone.0206009] [Medline: 30356273]
11. Cancela J, Koledova E, Restrepo M. Growth hormone treatment adherence in Latin American patients: Real-world data from the easypod Connect eHealth platform. XXVII Latin American Meeting of Pediatric Endocrinology. Cusco, Peru, October 24-27, 2018: Abstracts. Horm Res Paediatr 2018;90 Suppl 2:1-72 [FREE Full text] [doi: 10.1159/000493744] [Medline: 30347398]
12. Dunn TC, Xu Y, Hayter G, Ajjan RA. Real-world flash glucose monitoring patterns and associations between self-monitoring frequency and glycaemic measures: A European analysis of over 60 million glucose tests. Diabetes Res Clin Pract 2018 Mar;137:37-46 [FREE Full text] [doi: 10.1016/j.diabres.2017.12.015] [Medline: 29278709]
13. Calliari LEP, Krakauer M, Vianna AGD, Ram Y, Barbieri DE, Xu Y, et al. Real-world flash glucose monitoring in Brazil: can sensors make a difference in diabetes management in developing countries? Diabetol Metab Syndr 2020;12:3 [FREE Full text] [doi: 10.1186/s13098-019-0513-z] [Medline: 31921360]
14. Park S, Sentissi I, Gil SJ, Park W, Oh B, Son AR, et al. Medication Event Monitoring System for Infectious Tuberculosis Treatment in Morocco: A Retrospective Cohort Study. Int J Environ Res Public Health 2019 Jan 31;16(3):412 [FREE Full text] [doi: 10.3390/ijerph16030412] [Medline: 30709029]

XSL•FO
RenderX

15. Toscos T, Coupe A, Wagner S, Ahmed R, Roebuck A, Flanagan M, et al. Engaging Patients in Atrial Fibrillation Management via Digital Health Technology: The Impact of Tailored Messaging. J Innov Card Rhythm Manag 2020 Aug 15;11(8):4209-4217 [FREE Full text] [doi: 10.19102/icrm.2020.110802] [Medline: 32874747]
16. Toscos T, Drouin M, Pater JA, Flanagan M, Wagner S, Coupe A, et al. Medication adherence for atrial fibrillation patients: triangulating measures from a smart pill bottle, e-prescribing software, and patient communication through the electronic health record. JAMIA Open 2020 Apr 28;3(2):233-242 [FREE Full text] [doi: 10.1093/jamiaopen/ooaa007] [Medline: 32734164]
17. Koledova E, Tornincasa V, van Dommelen P. Analysis of real-world data on growth hormone therapy adherence using a connected injection device. BMC Med Inform Decis Mak 2020 Jul 29;20(1):176 [FREE Full text] [doi: 10.1186/s12911-020-01183-1] [Medline: 32727461]
18. Novillo-Ortiz D, Dumit EM, D'Agostino M, Becerra-Posada F, Kelley ET, Torrent-Sellens J, et al. Digital health in the Americas: advances and challenges in connected health. BMJ Innov 2018 Jul;4(3):123-127 [FREE Full text] [doi: 10.1136/bmjinnov-2017-000258] [Medline: 30101033]
19. Curioso WH. Building Capacity and Training for Digital Health: Challenges and Opportunities in Latin America. J Med Internet Res 2019 Dec 18;21(12):e16513 [FREE Full text] [doi: 10.2196/16513] [Medline: 31850849]
20. HealthIT.gov. Clinical Decision Support. 2018. URL: https://www.healthit.gov/topic/safety/clinical-decision-support [accessed 2021-03-01]
21. Lion F. Electronic Recording of Growth Hormone Dosing History: The Easypod™ Auto-Injector. CDTH 2010 Nov 01;5(4):271-276. [doi: 10.2174/157488510792927474]
22. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ 2007 Sep;85(9):660-667 [FREE Full text] [doi: 10.2471/blt.07.043497] [Medline: 18026621]
23. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl 2006 Apr;450:76-85. [doi: 10.1111/j.1651-2227.2006.tb02378.x] [Medline: 16817681]
24. Data Sharing Policy. Merck KGaA. URL: https://www.merckgroup.com/en/research/our-approach-to-research-and-development/healthcare/clinical-trials/commitment-responsible-data-sharing.html [accessed 2022-01-12]
25. Dimitri P, Fernandez-Luque L, Banerjee I, Bergadá I, Calliari LE, Dahlgren J, et al. An eHealth Framework for Managing Pediatric Growth Disorders and Growth Hormone Therapy. J Med Internet Res 2021 May 20;23(5):e27446 [FREE Full text] [doi: 10.2196/27446] [Medline: 34014174]
26. Lin CY, Updegraff JA, Pakpour AH. The relationship between the theory of planned behavior and medication adherence in patients with epilepsy. Epilepsy Behav 2016 Aug;61:231-236. [doi: 10.1016/j.yebeh.2016.05.030] [Medline: 27390026]
27. Cancela J, Guedes S, Koledova E. Real-world data from electronic monitoring of adherence to growth hormone treatment in children with growth disorders: a descriptive analysis. 2018 Presented at: The Endocrine Society's 100th Annual Meeting and Expo; Abstract no. A-6968; 2018; Chicago, IL, USA.
28. Bender BG, Ellison MC, Gleason M, Murphy JR, Sundstrom DA, Szefler SJ. Minimizing attrition in a long-term clinical trial of pediatric asthma. Ann Allergy Asthma Immunol 2003 Aug;91(2):168-176. [doi: 10.1016/S1081-1206(10)62173-4] [Medline: 12952111]
29. Cotter R, Burke J, Loeber R, Navratil J. Innovative retention methods in longitudinal research: A case study of the developmental trends study. J Child Fam Stud 2002;11(4):485-498. [doi: 10.1023/A:1020939626243]
30. Rodríguez-Arnao M, Sánchez A, López I, Fernández J, Bermúdez de la Vega J, Ballano V, et al. 55th Annual Meeting of the European Society for Paediatric Endocrinology (ESPE), Paris, France, September 10-12, 2016: Abstracts. Horm Res Paediatr 2016;86(Suppl 1):1-556 [FREE Full text] [doi: 10.1159/000449142] [Medline: 27592337]
31. Ayala-Estrada A, Antillon-Ferreira C, Saavedra-Castillo E, Barrientos-Pérez M, Rivero-Escalante H, Flores-Caloca O, et al. The Easypod Connect Observational Study (ECOS): Descriptive Analysis of Adherence to Treatment of Growth Hormone Deficient and Small for Gestational Age Naïve to Easypod Patients in Mexico 2012–2015. XXVI Annual Meeting of the Latin American Pediatric Endocrinology Society (SLEP). Buenos Aires, Argentina, November 8-11, 2016: Abstracts. Horm Res Paediatr 2016;86(Suppl 2):1-100 [FREE Full text] [doi: 10.1159/000451040] [Medline: 27816981]
32. Stoyanov G, Koledova E, Vandermeulen J, the Canadian ECOS group. Objectively measured treatment adherence in the easypod connect observational study (ECOS). Canadian interim analysis: population data and case reports. 2016 Presented at: Canadian Pediatric Endocrine Group 2016 Scientific Meeting; February 18-20, 2016; Edmonton, Alberta, Canada.
33. Malik S, Moloney C, Koledova E, Reston J, Weinman J. Designing a Personalized Digital Patient Support Program for Patients Treated With Growth Hormone: Key Design Considerations. J Med Internet Res 2020 Jul 29;22(7):e18157 [FREE Full text] [doi: 10.2196/18157] [Medline: 32723712]
34. Ng R, Carter SR, El-Den S. The impact of mobile applications on medication adherence: a systematic review. Transl Behav Med 2020 Dec 31;10(6):1419-1435. [doi: 10.1093/tbm/ibz125] [Medline: 31384950]
35. Peng Y, Wang H, Fang Q, Xie L, Shu L, Sun W, et al. Effectiveness of Mobile Applications on Medication Adherence in Adults with Chronic Diseases: A Systematic Review and Meta-Analysis. J Manag Care Spec Pharm 2020 Apr;26(4):550-561. [doi: 10.18553/jmcp.2020.26.4.550] [Medline: 32223596]

XSL•FO
RenderX

36. Pouls BPH, Vriezekolk JE, Bekker CL, Linn AJ, van Onzenoort HAW, Vervloet M, et al. Effect of Interactive eHealth Interventions on Improving Medication Adherence in Adults With Long-Term Medication: Systematic Review. J Med Internet Res 2021 Jan 08;23(1):e18901 [FREE Full text] [doi: 10.2196/18901] [Medline: 33416501]
37. Ferrante G, Licari A, Marseglia GL, La Grutta S. Digital health interventions in children with asthma. Clin Exp Allergy 2021 Feb;51(2):212-220 [FREE Full text] [doi: 10.1111/cea.13793] [Medline: 33238032]
38. Radovick S, Hershkovitz E, Kalisvaart A, Koning M, Paridaens K, Kamel Boulos MN. Gamification concepts to promote and maintain therapy adherence in children withgrowth hormone deficiency. J Multidiscip Sci 2018;1(1):71-80. [doi: 10.3390/j1010008]
39. Seiler P, Guimaraes F, Biglia L, Souza P. PDB14 - Follow-up of pediatric patients under treatment with somatropin in Brazil through a patient support program (PSP). Value in Health 2017 Oct;20(9):A921. [doi: 10.1016/j.jval.2017.08.2875]
40. Takeda A, Cooper K, Bird A, Baxter L, Frampton GK, Gospodarevskaya E, et al. Recombinant human growth hormone for the treatment of growth disorders in children: a systematic review and economic evaluation. Health Technol Assess 2010 Sep;14(42):1-209, iii-iv [FREE Full text] [doi: 10.3310/hta14420] [Medline: 20849734]

Abbreviations

ΔHSDS: change in height standard deviation score
ADSS: adherence decision support system
ECOS: Easypod Connect Observational Study
EHR: electronic health record
GH: growth hormone
HCP: health care provider
HSDS: height standard deviation score
PSP: patient support program
rhGH: recombinant human growth hormone

Edited by G Eysenbach; submitted 04.08.21; peer-reviewed by R Ng; comments to author 14.09.21; revised version received 21.10.21; accepted 27.11.21; published 20.01.22.

Please cite as:
Assefi A, van Dommelen P, Arnaud L, Otero C, Fernandez-Luque L, Koledova E, Calliari LE
Adherence to Growth Hormone Treatment Using a Connected Device in Latin America: Real-World Exploratory Descriptive Analysis Study
JMIR Mhealth Uhealth 2022;10(1):e32626
URL: https://mhealth.jmir.org/2022/1/e32626
doi:10.2196/32626
PMID:35049518

©Aria Assefi, Paula van Dommelen, Lilian Arnaud, Carlos Otero, Luis Fernandez-Luque, Ekaterina Koledova, Luis Eduardo Calliari. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 20.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Attitudes Toward Mobile Apps for Pandemic Research Among Smartphone Users in Germany: National Survey

Lorina Buhr[1*], MA; Silke Schicktanz[1*], PhD, Prof Dr; Eike Nordmeyer[1,2], MSc

[1]Department of Medical Ethics and History of Medicine, University Medical Center Göttingen, Göttingen, Germany

[2]Department of Agricultural Economics and Rural Development, University of Göttingen, Göttingen, Germany

[*]these authors contributed equally

Corresponding Author:
Silke Schicktanz, PhD, Prof Dr
Department of Medical Ethics and History of Medicine
University Medical Center Göttingen
Humboldtallee 36
Göttingen, 37073
Germany
Phone: 49 5513969009
Email: sschick@gwdg.de

Abstract

Background: During the COVID-19 pandemic, but also in the context of previous epidemic diseases, mobile apps for smartphones were developed with different goals and functions, such as digital contact tracing, test management, symptom monitoring, quarantine compliance, and epidemiological and public health research.

Objective: The aim of this study was to explore the potential for the acceptance of research-orientated apps (ROAs) in the German population. To this end, we identified distinctive attitudes toward pandemic apps and data sharing for research purposes among smartphone users in general and with a focus on differences in attitudes between app users and nonusers in particular.

Methods: We conducted a cross-sectional, national, telephone-based survey of 1003 adults in Germany, of which 924 were useable for statistical analysis. The 17-item survey assessed current usage of pandemic apps, motivations for using or not using pandemic apps, trust in app distributors and attitudes toward data handling (data storage and transmission), willingness to share coded data with researchers using a pandemic app, social attitudes toward app use, and demographic and personal characteristics.

Results: A vast majority stated that they used a smartphone (778/924, 84.2%), but less than half of the smartphone users stated that they used a pandemic app (326/778, 41.9%). The study focused on the subsample of smartphone users. Interestingly, when asked about preferred organizations for data storage and app distribution, trust in governmental (federal or state government, regional health office), public-appointed (statutory health insurance), or government-funded organizations (research institutes) was much higher than in private organizations (private research institutions, clinics, health insurances, information technology [IT] companies). Having a university degree significantly ($P<.001$) increased the likelihood of using a pandemic app, while having a migration background significantly ($P<.001$) decreased it. The overwhelming majority (653/778, 83.9%) of smartphone users were willing to provide their app data for state-funded research. Regarding attitudes toward app usage, striking differences between users and nonusers were found. Almost all app users (317/327, 96.9%) stated they would be willing to share data, whereas only 74.3% (336/452) of nonusers supported data sharing via an app. Two-thirds (216/326, 66.3%) of app users fully or rather *agreed* with the statement that using a pandemic app is a social duty, whereas almost the same proportion of nonusers entirely or rather *disagreed* with that statement (273/451, 60.5%).

Conclusions: These findings indicate a high potential for the adoption of ROAs among smartphone users in Germany as long as organizational providers engaged in development, operation, and distribution are state-funded or governmental institutions and transparency about data-using research institutions is provided.

(JMIR Mhealth Uhealth 2022;10(1):e31857) doi:10.2196/31857

KEYWORDS

user; pandemic; smartphone apps; mobile apps; telephone-based survey; Germany; data sharing; data donation; ethics; trust; COVID-19; mHealth; mobile applications; digital health; health applications

XSL•FO
RenderX

Introduction

Background

After the outbreak of the COVID-19 pandemic, various governments and the European Union decided that digital solutions, most notably smartphone apps, should contribute to pandemic management and research [1,2]. Four different digital public health technologies have been described: (1) mobile apps for proximity and contact tracing, (2) mobile and web apps for symptom monitoring, (3) digital tools for quarantine compliance, (4) data analytic tools for flow modelling [3]. In particular, digital contact tracing has then been intensively debated from practical and ethical perspectives [4-9]. Recently, digital options for providing proof of individual immunization or health status, such as the "Digital Green Certificate" proposed by the European Commission [10], and "check-in-apps," such as the German "Luca-App," have also been the subject of contentious debate. In contrast, apps with a primary function of transferring or making available digital data to pandemic-related epidemiological and public health research have been far less publicly discussed. Some of these apps can be connected with a wearable device such as a fitness watch or fitness tracker. We call this type of app a research-orientated app (ROA). ROAs promise to provide answers to various research questions in the field of (digital) epidemiology and public health research. The German Data Donation App [11] is a classic example of an ROA-type pandemic app. From an ethical and social perspective, however, various issues need to be addressed: the consent to collect and the protection of sensitive data (eg, indicating bodily activity and movement); governance structures for data sharing and usage; and public support or even social obligations for such digital health technologies in pandemic research [8,12]. States that have comparatively strict data protection laws and those that are defending permissive standards regarding citizens' rights such as China, South Korea, or Israel have to find different solutions for consent and voluntary data sharing, whether for pandemic management or research.

Previous Work

In recent years, there has been an increasing amount of qualitative, quantitative, and mixed methods studies in an international context that have explored behaviors, motivations, and perceptions about mobile phone–based apps for health. They focus on health apps in general [13,14] or on health apps for a specific field of disease (eg, chronic diseases) [15,16]. As pandemics apps are a type of health app, studies on pandemic apps can be considered a further type of domain-specific health app. Since the start of the COVID-19 pandemic, most research on pandemic apps has been carried out on contact tracing apps, mainly focusing on perceptions toward digital contact tracing apps as well as motivations, acceptability, drivers, and barriers for app uptake [17-23]. Most of these studies have focused on privacy and surveillance concerns, including questions of trust and mistrust in different app providers. In most cases, these are national studies conducted as online panels. An exception is the Ipsos Mori survey in the United Kingdom, which was also telephone-based [24]. There are already some cross-national studies that have surveyed the conditions for acceptance of contact tracing apps [25-30]. In the German-speaking context, the studies by Becker et al [20], Kaspar [31], Buder et al [32], and the eGovernment Monitor [33] should be highlighted, since they were published before our data collection started and gave input for questionnaire construction in this study. Studies in the German-speaking context that mainly focused on privacy and surveillance concerns reported a relatively high rate of people who doubt the fundamental benefit of contact tracing apps: around one-third and up to one-half of the respondents were skeptical about using an app [18,33]. Moreover, results differed as to which organizations and providers are trusted in connection with the development and release of smartphone apps and to what extent.

In the context of health apps in general and pandemic apps in particular, current debate is mainly focusing on privacy concerns and perceptions toward sharing health data [13,34]. Beierle et al [35] found that there is a complex picture to describe smartphone users' willingness to share data with researchers, showing that privacy concerns are not clearly the main factor for not permitting data sharing; personality traits, gender, and age are also considerable factors. Kaspar [31] provided a valuable multiple regression analysis that indicated significant differences in motivations using a contact tracing app or the German Data Donation App (n=406, convenience sample). Interestingly, he found that "motivation for providing the personal data requested by the individual app type was also higher in the case of the contact tracing app (mean 4.48, SD 2.32) compared to the Data Donation app (mean 3.41, SD 2.23; t_{405}=10.86, P<.001, d=0.54)" [31]. Recently, von Wyl et al [22] reported results from their nationwide online survey panel in Switzerland describing differences between users and nonusers of pandemic apps. To the best of our knowledge, there is no study that has systematically analyzed differences between app users and nonusers of pandemic-related apps for German smartphone users with regard to ROA.

Objectives

This study explored the potential for the adoption of ROA among the German population. We focused on smartphone users and aimed to identify specific challenges for app usage. To our knowledge, this is the first nationwide, telephone-based survey study in Germany since the first pandemic apps ("Corona-Warn-App" and "Data-Donation-Apps") were released nationally. It is also, to our knowledge, the first study focusing primarily on individual data sharing via smartphone apps for pandemic research. Leading research questions were (1) "Which sociodemographic and personal factors influence the use of a pandemic app among smartphone users in the German population?" and (2) "How do users and nonusers of pandemic apps differ in their motives, attitudes toward pandemic apps, and willingness to share data with researchers?" The objectives of this paper were therefore to identify distinctive attitudes toward pandemic apps and data sharing for research among smartphone users in general and with a focus on differences in attitudes between app users and nonusers in particular. The results can inform empirically based ethical recommendations for the future development, design, and implementation of ROA.

XSL•FO
RenderX

Methods

Study Design

We designed a survey comprised of 17 question units (see Multimedia Appendix 1) to explore attitudes toward pandemic apps and toward data sharing for research. The survey study was approved by the local Human Research Review Committee (Reference Number 4/12/20) at the University Medical Center Göttingen.

A representative phone-based survey seemed more appropriate than online panels to reach people who are not internet-savvy or avoid online surveys. Especially when it comes to questions of public acceptance of modern technologies, such as smartphone apps, a broader sampling strategy seemed more appropriate to make statements about the whole population.

Inclusion criteria were people (1) aged ≥18 years, (2) with a registered address in Germany, and (3) who were literate in German. The sample population was comprised of private households in Germany with at least one landline connection and people with at least one mobile phone connection. The population survey was conducted by the company Kantar GmbH. It took approximately 15 minutes to 20 minutes to complete the questionnaire. A dual-frame sampling approach (ie, taking into account both landline and mobile phone numbers) was used. The landline (n=703) and cell phone (n=300) samples were then combined by statistical weighting according to the demographic statistics of the German population (see Multimedia Appendix 2). The survey was conducted anonymously, hence no identifying data were included in the data file Kantar GmbH sent to the authors. Due to incomplete data, 79 cases were excluded from the sample. Thus, our population sample included 924 cases, which is also the number of complete interviews.

Sample

A representative telephone-based population survey with 1003 people in Germany aged 18 years or older was conducted between December 10, 2020 and January 18, 2021. A sample size of 1000 interviews was originally planned; 3 further interviews were conducted due to already arranged appointments with target persons. The German population aged ≥18 years currently is around 69.4 million. In current survey research, 1000 respondents have proven to be a practicable and statistically acceptable sample size for representative population surveys in Germany. We can refer to the seminal national survey study by Richter et al [36] on acceptability of data donation in medical and health contexts in Germany, which also included a similar sample size (n=1006) using the population-representative survey panel “forsa.Omninet” by the German Forsa Institute [37]. See Multimedia Appendix 3.

Survey Items

The survey instrument for the phone questionnaire contained closed-ended question types. The questionnaire encompassed 17 question units in the German language and entailed the following domains: (1) current usage of smartphone and pandemic apps, (2) motivations for using or not using pandemic apps, (3) trust in app distributors and data storage, (4) willingness to share coded data with research institutions using a pandemic app and attitudes toward data handling, (5) social attitude toward app use, and (6) demographic and personal characteristics (Multimedia Appendix 1). The composition was informed by an analysis of then-existing surveys about pandemic apps and of another survey we conducted in 2020 on attitudes toward data sharing of wearable data among cardiac patients (publication in preparation). “Pandemic apps” were defined in this survey as native mobile applications for smartphones specifically designed for the containment, management, and research of epidemic and pandemic infectious diseases. Since a pilot test showed that there was no broad understanding of different types of app construction—standalone apps and web apps—we limited the definition to standalone smartphone apps (cf, a study on pandemic web apps by Scherr et al [38]). Multiple answers were possible for 6 question units, 5 questions were formulated as yes/no questions (“yes,” “no,” “I don't know,” “other reason”), and 2 items contained a Likert scale: 1 item with a 4-point Likert scale and the other with a 5-point Likert scale. If a response other than the given answer choices was given, this was recorded unaided by the interviewers. Questions were presented to each participant in the same order; however, the order of within-item responses was randomly assigned to reduce response-set bias. Kantar GmbH was responsible for the implementation of the questionnaire for fieldwork (eg, programming, codes, filter guidance) and pretesting with regard to understandability.

Statistical Analysis

Descriptive statistics were calculated for all items. Since we were interested in pandemic app usage and attitudes toward data sharing with research institutions, we focused on the subsample of smartphone users in our statistical analysis (n=778). Chi-squared tests were used to identify differences between users and nonusers of the app regarding the willingness for data sharing, social attitudes toward app usage, perceptions of the trustworthiness of the app provider, and the preferred location for data storage. To quantify the factors influencing app usage, logistic regression analysis was used. For this purpose, sociodemographic variables such as age, gender, education, place of residence, and migration background were included in the model. In addition, personal experience of being directly infected or knowing someone who has been infected by the COVID-19 virus was elicited (personal affection). Statistical significance was determined by P values <.05. Likert-scale answers were pooled into categories (eg, “fully agree” and “rather agree” into “fully/rather agree” and “entirely not agree” and “rather not agree” into “entirely not/rather not agree”). All statistical analyses were carried out using SPSS version 26 (IBM Corp, Armonk, NY). Due to the fact that we used weighted data, the sample size may differ by ±1 in some analyses due to rounding effects. Calculating with weighted data also has the effect that percentages can deviate minimally in the decimal place compared with the quotient n/N in natural numbers. We marked all cases in which this deviation occurred with an asterisk (*) or respectively, a reference mark in tables. For a detailed description of the weighting, see Multimedia Appendix 2. Logistic regression analysis was conducted, as the statistical assumptions were met, and all rating scales were treated as

XSL•FO
RenderX

nominal (except gender as an ordinal variable), including the dependent variables. Ordinal regression models, however, could be considered as an alternative approach (with distinct limitations).

Results

Demographic Characteristics of Smartphone Users

Of the 924 participants that were included for statistical analysis, 84.2% (778/924) stated that they used a smartphone. The following analysis refers to this subset of smartphone users, as we deemed smartphone usage a condition for technology-specific considerations on pandemic app usage. We assumed that inclusion of non-smartphone users in the statistical analysis would have engendered a mixed sample of distinctive versus hypothetical usage attitudes.

Table 1 presents the demographic and personal characteristics of the sample of the survey participants and the subsample of smartphone users. Compared with the whole sample, whose demographic characteristics are representative of the German population, smartphone users differed slightly in 3 regards: (1) They were younger, (2) they were more likely to have a higher level of education, and (3) they were personally affected by the COVID-19 pandemic slightly more often (affectedness was reported according to their statements). The mean age of the survey sample was 50.19 (SD 17.96) years, and age ranged from 18 years to 95 years; the mean age of the analytic sample was 46.84 (SD 16.96) years, and age ranged from 18 years to 95 years. A total of 37.2%* (290/778) of the smartphone users indicated that they had an A-level or university degree; this is slightly more than the participant sample of whom only 32.8%* (304/924) reported having an A-level or university degree. In terms of COVID-19 infection, 37.8% (294/778) of the smartphone users reported they had been personally affected, whereas 34.1% (315/924) of survey participants reported being personally affected.

XSL•FO
RenderX

Table 1. Sociodemographic profile of survey participants (n=924) and the subsample of smartphone users (n=778).

Characteristic	Survey participants (n=924), n (%)	Smartphone users (n=778), n (%)
Gender		
Female	474 (51.3)	404 (51.9)
Male	450 (48.7)	374 (48.1)
Nonbinary	0 (0)	0 (0)
Age (years)		
18-30	153 (16.6)	153 (19.7)
30-39	120 (13.0)	119 (15.3)
40-49	175 (18.9)	168 (21.5[a])
50-59	183 (19.9[a])	153 (19.6[a])
60-69	128 (13.8[a])	94 (12.1)
≥70	165 (17.9)	91 (11.7)
Education		
None/still in school	51 (5.5)	44 (5.7)
Without A-level	570 (61.7)	444 (57.1)
A-level	138 (14.9)	133 (17.1)
Academic degree	166 (17.9[a])	157 (20.1[a])
Immigration background		
Yes	170 (18.4)	156 (20.1)
No	754 (81.6)	622 (79.9[a])
Region		
West German states	765 (82.8)	645 (83.0[a])
East German states (including Berlin)	159 (17.2)	133 (17.0[a])
Personally affected by COVID-19 infection or knowing someone who was		
Yes	315 (34.1)	294 (37.8)
No	609 (65.9)	484 (62.2)

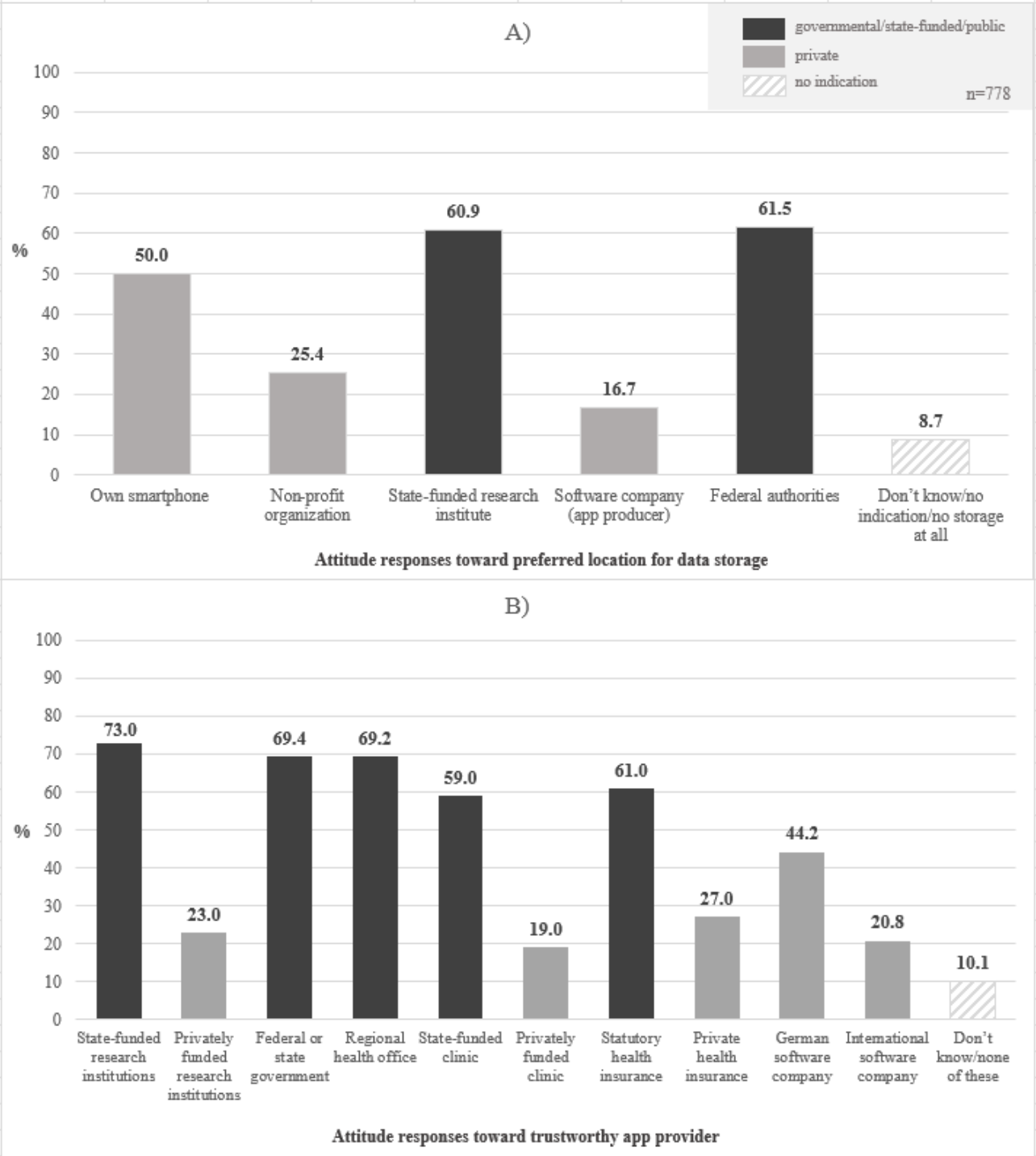
[a]Due to the fact that we used weighted data, the sample size may differ by ±1 in some analyses due to rounding effects. Calculating with weighted data also has the effect that percentages can deviate minimally in the decimal place compared with the quotient n/N in natural numbers. For a detailed description of the weighting, see Multimedia Appendix 2.

Attitudes Among Smartphone Users Toward Pandemic App Providers and Toward Data Sharing With Research Institutes

Our analysis of attitudes focused on 2 major topics. First, we report attitudes toward app providers, by which we understand organizations and institutions involved in the development, provision, and operation of pandemic apps (see Figure 1). Second, we present results on attitudes toward sharing data collected by a pandemic app for research (Table 2, Table 3, and Multimedia Appendix 4). The descriptive analysis on attitudes is supplemented by a regression analysis on app usage among smartphone users (Table 4). In a third step, we focused on differences between users and nonusers of pandemic apps on these and related issues (Figures 2 and 3).

XSL•FO
RenderX

Figure 1. Attitude responses toward (A) preferred location for data storage and (B) trustworthy app provider among smartphone users (n=778).



XSL•FO
RenderX

Table 2. Attitudes among people willing to share data ("data sharers," n=653) for research via an app among smartphone users (n=778).

Attitude responses among data sharers (n=653)	Results, n (%)
Kind of data to be shared[a]	
Don't know/none of these	10 (1.5)
Data collected by a fitness watch	215 (33.0[b])
Continuous data (ambient temperature)	298 (45.6)
Health-related data	330 (50.5)
Location and movement data	366 (56.1[b])
Contacts with other people	436 (66.8)
Data manually entered in the app	447 (68.5)
Test results	553 (84.8[b])
Preferred way and mode of data transmission to the research institute[a]	
Don't know/none of these	8 (1.2)
Calling a video hotline of the research institute	58 (8.9)
Calling a telephone hotline of the research institute	101 (15.5)
Sending the data via SMS	123 (18.9[b])
Sending the data via email	169 (25.9)
By entering the data on the website of the research institute	207 (31.7)
Sending the data automatically to the research institute	379 (58.1[b])
By enabling data sharing in the app each time	437 (66.9)
Transparency about data-using research institutes	
Not so important/not at all important	158 (24.2)
Very/rather important	495 (75.8)

[a]Multiple answers were possible.

[b]Due to the fact that we used weighted data, the sample size may differ by ±1 in some analyses due to rounding effects. Calculating with weighted data also has the effect that percentages can deviate minimally in the decimal place compared with the quotient n/N in natural numbers. For a detailed description of the weighting, see Multimedia Appendix 2.

Table 3. Attitudes among people not willing to share data ("non-data sharers," n=125) for research via an app among smartphone users (n=778).

Attitude responses among non-data sharers (n=125)	Results, n (%)
Why people do not share data[a]	
Don't know	3 (2.2[b])
Other reasons	9 (7.2)
I am worried that the data will be leaked.	78 (62.2[b])
I doubt that this data will help research.	78 (62.4)
I am concerned about unknown third parties using my data.	85 (68.2[b])

[a]Multiple answers were possible.

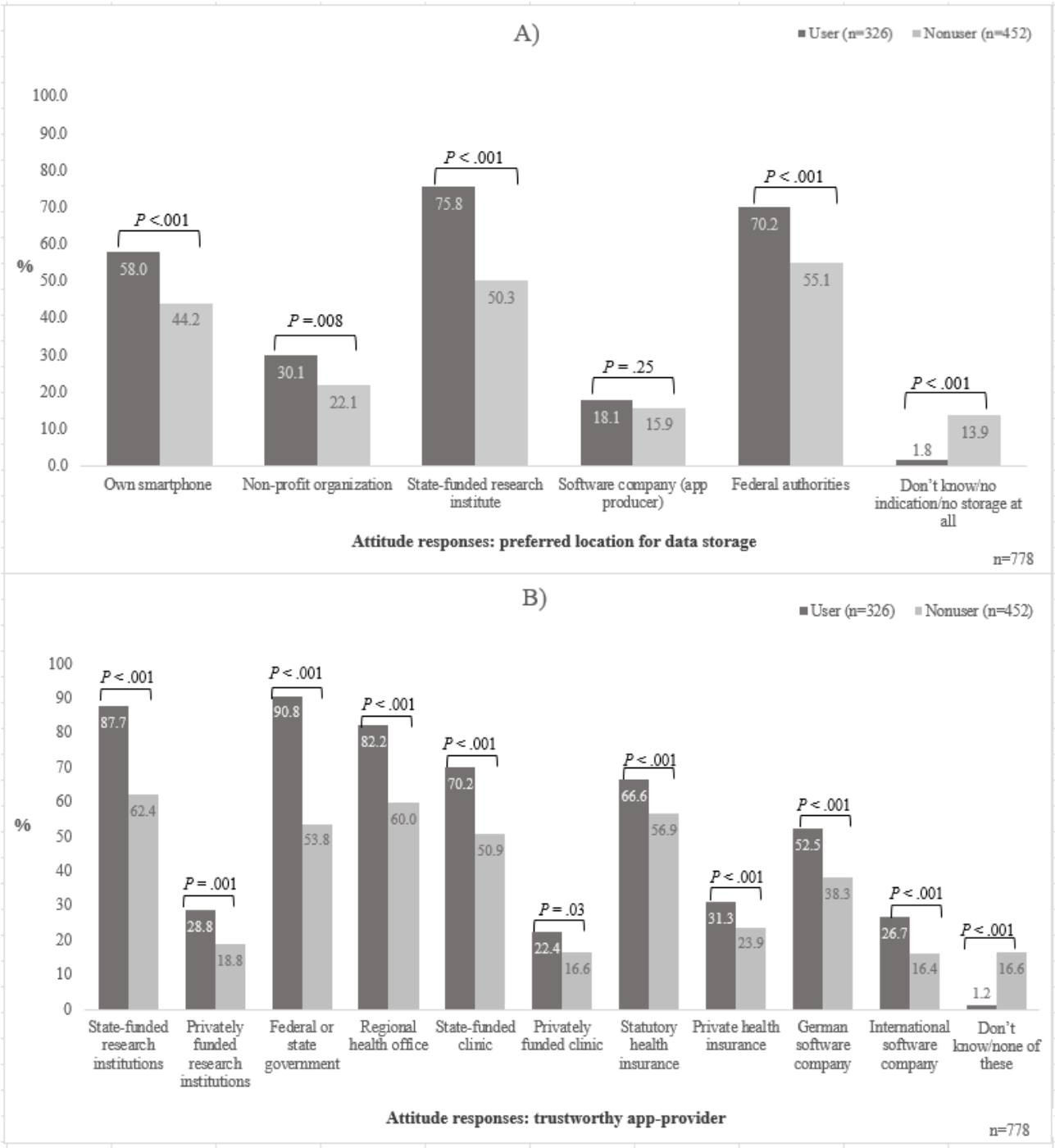
[b]Due to the fact that we used weighted data, the sample size may differ by ±1 in some analyses due to rounding effects. Calculating with weighted data also has the effect that percentages can deviate minimally in the decimal place compared with the quotient n/N in natural numbers. For a detailed description of the weighting, see Multimedia Appendix 2.

XSL•FO
RenderX

Table 4. Multivariable correlates of pandemic app usage.

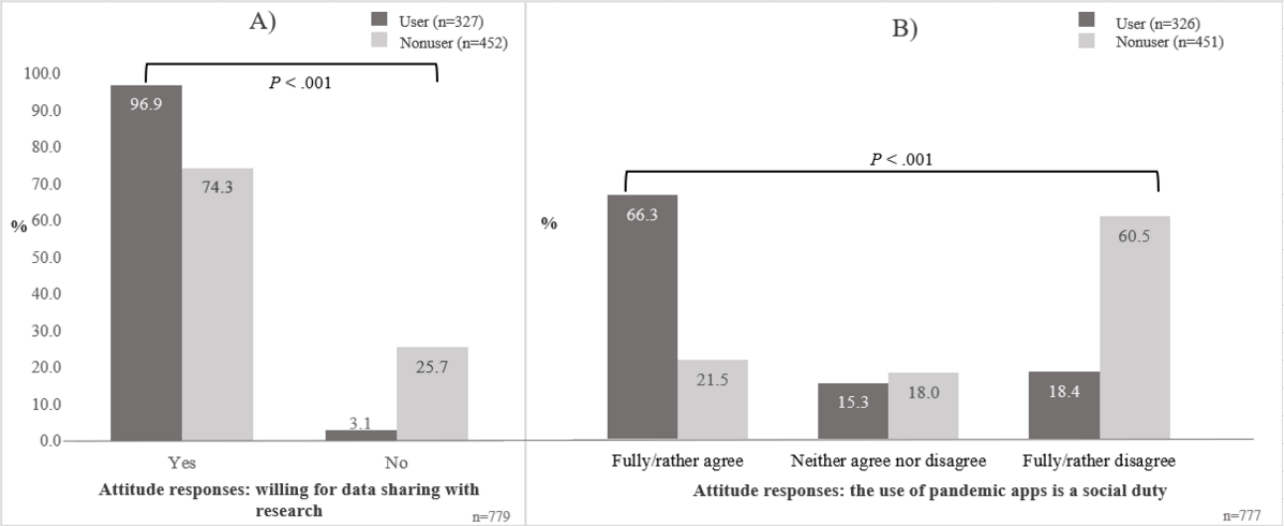
Variable	Coefficient	*P* value
Constant	–0.564	.02
Male (vs female)	0.069	.65
Age groups	0.055	.27
University degree (vs no university degree)	1.081	<.001
Eastern Germany (vs western Germany)	–0.494	.02
Immigration background (vs no immigration background)	–1.242	<.001
Being affected (vs not personally affected)	0.279	.09

Figure 2. Different attitude responses toward (A) preferred location for data storage and (B) trustworthy app provider among app users (n=326) and nonusers of pandemic apps (n=452).



XSL•FO
RenderX

Figure 3. Different attitude responses toward (A) data sharing for research and (B) the statement, "the use of pandemic apps is a social duty" among app users (n=327 and n=326, respectively) and nonusers of pandemic apps (n=452 and n=451, respectively). Due to the fact that we used weighted data, the sample size may differ by ±1 in some analyses due to rounding effects. For a detailed description of the weighting, see Multimedia Appendix 2.



Attitudes Toward App Providers: Trust in State (-Funded) Organizations

Our survey revealed that German smartphone users demonstrated a strong preference for state-funded or governmental organizations with regard to storage of app data as well as app providers when asked which providers the participants considered most trustworthy. We considered the statement of preference for a certain app actor for data storage to be an indication of trust in that app actor (or one's smartphone) for data storage and management. As Figure 1A indicates, the majority of smartphone users preferred state-funded research institutes (474/778, 60.9%) or federal authorities (478/778, 61.5%*) when asked where the data for the pandemic app should be stored. It is noteworthy that those rates are even higher than those for using one's own smartphone as data storage (389/778, 50.0%). Among smartphone users, the least preferred storage location was the software company producing the pandemic app (130/778, 16.7%). This pattern has been consistently mirrored when it comes to trust attitudes toward distributors of pandemic apps. Trust in governmental (federal or state government, regional health office), public-appointed (statutory health insurance), or government-funded research institutes or organizations as app distributors was much higher than trust in private organizations (research institutions, clinics, health insurance). For example, 61.0%* (475/778) of participants considered public health insurance trustworthy distributors for pandemic apps, but only 27.0% (210/778) reported the same for private health insurance.

Interestingly, German software companies were classified twice as trustworthy (344/778, 44.2%) as international companies (161/778, 20.8%*). Figures 1A and 1B illustrate the obvious differences in trust between state(-funded) and private organizations.

Factors Influencing the Usage of a Pandemic App

Less than half of participants stated that they used a pandemic app (326/778, 41.9%).

We used logistic regression analysis to determine which sociodemographic characteristics influenced the probability of using or not using a pandemic app. Having a university degree significantly (P<.001) increased the likelihood of using a pandemic app, while having a immigration background significantly (P<.001) decreased the likelihood of using a pandemic app (see Table 4). Furthermore, residence in the eastern part of Germany reduced the likelihood of using a pandemic app compared with a residence in the western part (P=.02; Multimedia Appendix 5). No significant influence was shown for age, gender, or being personally affected by COVID-19 or having relatives or friends who have been infected.

Attitudes Toward Data Sharing With Research: High Willingness

The high level of trust in state-funded app providers is matched by the fact that 83.9% (653/778) of smartphone users were willing to provide their app data for state-funded research. We called that subsample of people willing to share data with research institutions "data sharers." When asked about what kind of data people were willing to share for research, the vast majority indicated test results (553/653, 84.8%*), followed by contact tracing data (436/653, 66.8%). In contrast, only one-third (215/653, 33.0%*) were willing to share data via a pandemic app that were originally collected by a digital mobile device, such as a fitness watch (Table 2, Multimedia Appendix 4). Furthermore, there was a distinct preference for entering data manually into an app (447/653, 68.5%) versus data collected continuously and automatically (298/653, 45.6%). The preferred mode of data transmission, however, was not uniform: Two-thirds (437/653, 66.9%) were in favor of user-initiated data transmission, whereas a large number (379/653, 58.1%*) of respondents affirmed automatic data transmission to a research institute. For one-third of data sharers (207/653, 31.7%), manual entry on a research institute's website was also an option. Notably, there was a clear trend when it came to the importance of transparency about which research institutions will use the app data. For the vast majority of data sharers

XSL•FO
RenderX

(495/653, 75.8%), it was very or rather important that those research institutes were clearly designated (Table 2, Multimedia Appendix 4). Overall, these results indicated a high willingness for data sharing for public research. Interestingly, among the data sharers, a majority expressed a wish for self-controlled data transmission (437/653, 66.9%) and transparency about the involved data-analyzing institutions (495/653, 75.8%).

The 3 most frequent reasons why people were not willing to share their data for research (125/778, 16.1%) were concerns about lack of control of app data (85/125, 68.2%*; ie, the concern that third parties were using the data without consent), that data would be leaked (78/125, 62.2%*), and doubts on whether app data really would bring research forward (78/125, 62.4%; Table 3, Multimedia Appendix 4).

Nonusers of Pandemic Apps Have Less Trust in State-Funded Organizations

We found substantial differences in attitudes toward pandemic app providers between users and nonusers of pandemic apps (Figure 2). Nonusers of pandemic apps showed higher approval for public government or state-funded organizations but their trust regarding data storage and distribution of pandemic apps is considerably lower. Figures 2A and 2B provide a comparison of pandemic app users versus nonusers. Although 75.8% (247/326) of pandemic app users preferred state-funded research institutes for data storage, only 50.3% (227/451) of nonusers did so ($P<.001$). Furthermore, 90.8% (296/326) of pandemic app users indicated trust in federal and state governments as app providers, compared with only 53.8% (243/452) of nonusers ($P<.001$). However, no statistically significant differences between users and nonusers of pandemic apps were found regarding software companies as a preferred location for data storage ($P=.25$).

Nonusers of Pandemic Apps Are Less Willing to Share Data With Research Institutes

As indicated in Figure 3A, there was high covariance between app usage and the willingness to share data with research institutes via a pandemic app. The overwhelming majority (317/327, 96.9%) of app users stated they would be willing to share data, whereas only 74.3% (336/452) of nonusers supported data sharing via an app; thus, users and nonusers differed significantly in their attitude toward data sharing for state-funded research via an app ($P<.001$). This is consistent with our findings on decreased trust in state-funded research institutes as data storage and app providers (see the previous section). The differences between users and nonusers of pandemic apps were most apparent in the moral and social attitudes toward pandemic app usage (Figure 3B). Two-thirds of app users (216/326, 66.3%) fully or rather *agreed* with the statement that using a pandemic app is a social duty, whereas almost the same proportion of nonusers completely or rather *disagreed* with that statement (273/451, 60.5%). Thus, the moral and social attitude for app usage was inverted between users and nonusers of pandemic apps.

Discussion

Principal Findings and Comparison With Prior Work

Overview

This study examined pandemic app usage and attitudes toward data sharing with research institutes among a sample of smartphone users, which represented a subsample of a representative population survey in Germany. Our study provides several important findings. In the following sections, we focus on 4 of them. First, the results showed a high willingness to share data with state-funded research institutes for pandemic research, but the willingness for data sharing went along with a strong need for self-controlled data handling and transparency about the involved data-analyzing research institutions. Second, there was a remarkable decline in trust toward private providers and organizations involved in data storage and distributing pandemic apps when compared with state-funded organizations. Third, regression analysis showed app usage is positively correlated with a higher level of education. Fourth, our study revealed significant differences in trust attitudes between app users and nonusers.

High Willingness for "Self-Determined" Data Sharing With Research Institutes

One of the aims of this study was to elicit peoples' attitudes and concerns toward sharing data collected by pandemic apps. The overwhelming support for data sharing via pandemic apps for research purposes among smartphones users in Germany is consistent with previous surveys that examined the willingness for "data sharing" or "data donation" [39]. For example, the online survey panel by tmf/Medical Informatic Initiative and Richter et al [36] indicated a consent rate of 78.8% (n=1006) for "data donation" (ie, a consent-free approach) for medical research among German adults in 2019 (cf, Mello et al [40] for patients' views on data sharing with research in the United Kingdom). Becker et al [20] showed that a large portion of participants disagreed with providing governmental organizations with anonymous user data to contain the pandemic in Germany. However, the results of the study by Becker et al [20] showed that app adoption was not negatively affected if the data-receiving organizations were public health authorities or research institutes.

In the international context, the picture on the willingness to share data in general is quite heterogeneous. For example, Abeler et al [41] reported that 64.6% of survey participants in the United Kingdom (n=1055) would permit data sharing with researchers after the pandemic. Maytin et al [23] reported for young adults in the United States that 45.1% (231/513) would agree or strongly agree with actively providing health data via an app. A possible explanation for this might be that answers on this topic strongly depend on how the question is posed, how the app providers are presented, and which data participants are asked to share for what purpose.

Implications for Policy Makers

When informing policy makers about affirmative attitudes among smartphone users toward willingness to share data with

XSL•FO
RenderX

research institutes, those findings should be contextualized with stated preferences for criteria of data sharing and processing. In our study, we explored different criteria that allow for a more nuanced picture: (1) the kind of data to be shared (eg, health data, location data), (2) transparency about data-receiving research institutions, (3) the mode of data collection, and (4) data transmission (Table 2). Our findings indicated that, to gain sufficient rates of people sharing app data for pandemic research purposes, pandemic apps should have the following features: (1) type of data to include test results, contacts with persons, location, and movement data (less support for data from a fitness watch); (2) detailed transparency about data-receiving research institutes (vs no transparency about the data receiver); (3) manually entered data (vs automatically collected data); (4) manually enabled data transmission *or* automatic sending; (5) storage of collected or disclosed data via pandemic app on servers of the respective state-funded research institute (least support for app storage by private organizations such as tech companies).

In summary, we conclude that those willing to share data for research purposes express a strong interest in a self-determined way of data sharing. This means that mechanisms of manual handling such as activation of data transfer, a set of selected kinds of data, and comprehensible and detailed information about data processing would likely increase willingness for data sharing with research institutions. However, since automatic data transmission is also endorsed by a large portion of participants, the picture is more complex. An ambiguous tendency in attitudes toward data transmission was also reported by Becker et al [20]. Hence, an option might be to provide app users with the option to select between automatic and manual data transmission. Further research is then needed to examine users' long-term satisfaction with these options. The need for more research in this area is also reinforced by recent studies that indicate that the preferred kinds of data willing to be shared may also differ among age groups [23]. Furthermore, the issue of a self-determined manner of data sharing should be examined in relation to "eHealth literacy," sometimes also called "media health literacy" or "(digital health) data literacy" [42,43]. To measure eHealth literacy in the context of pandemics, the eHealth Literacy Scale (eHEALS) developed by Norman and Skinner [42,44] seems very promising. Future research should test whether eHealth literacy positively correlates with beliefs in the benefits of (pandemic) apps and the willingness to share data with pandemic research (see, for example, the patient survey study by Knitza et al [45] in rheumatology using the validated German version of eHEALS [46]).

Gap in Trust Between State-Funded and Private Organizations

One important finding in our regard is the extent to which attitudes toward state-funded and private app providers vary among smartphone users: Private providers were considerably less trusted with data storage and providing an app. Here, we interpreted the preference of a storage location as an expression of trust in this specific organization. We found that trust in state-funded research institutes and governments for the storage of app data is very high (almost two-thirds of smartphone users). This is an encouraging message for state-funded research intuitions even if ongoing public discussions about privacy, data security, governmental surveillance practices, and centralized versus decentralized storage solutions for pandemic apps might give the opposite impression [17,47,48]. However, there is paradox-like situation. On the one hand, current debates on privacy and the willingness of governmental providers to take those concerns into account (eg, as with the German Corona-Warn-App, using open-source code and decentralized storage) might have been seen as trust-building efforts in favor of democratic governments and as a clear demarcation from state surveillance tendencies. On the other hand, increasing awareness and media reports on how information technology (IT) companies use data streams and cloud backups (eg, Apple iCloud or Google Cloud) may have also increased skepticism toward such providers, especially when it comes to public goods, such as health issues. National or cross-country surveys such as those by Simko et al [25], Hargittai et al [28], and Wiertz et al [21] as well as the British IPSOS MORI report [24] support this interpretation. They also report a disparity in terms of trustworthiness between government agencies, health departments, and IT companies, either big tech companies or start-ups. However, there are noticeable national differences. For example, the studies by Wietz et al [21] in the United Kingdom and Hargittai et al [28] in the United States reported a significant difference in public trust between the national government and the top national health authorities (National Health Service in the United Kingdom and Health Protection Agency in the United States), with the latter being significantly more trusted. In contrast, our results cannot confirm this kind of split concerning trust in official health authorities and research institutions in Germany. Although for Kaspar [31], it was still an open question in 2020 "as to whether different providers are assessed as having different levels of trust" [31], our findings provide a clear answer to this question.

The Challenge for Public-Private Partnerships for Pandemic Apps

The large gap between state-funded and private providers poses a challenge for the reality of pandemic app development, which is mainly achieved via public-private partnerships. Considering that pandemic research is of extremely high public health relevance and therefore differs from many (not all) other areas of mobile health (mHealth) where health behavior or health research addresses a smaller population of patients, research on pandemic apps can clearly benefit from a strong emphasis on the public partner. However, the high level of trust in government and state-funded research institutes as app providers can be gambled away if there is an increasing reliance on private-public partnerships in which tech companies co-determine the technical and design solutions, as was the case when Apple and Google offered governments their common exposure notification application programming interface (API) [4,49,50]. However, digital contact tracing apps in the COVID-19 pandemic have not yet a reached a sufficient level of broad uptake, such as at least 60% to 70%, which in turn is necessary for validating them as effective tools for pandemic containment and management [51,52]. In our case, 41.9% (326/778) of participants confirmed the use of a pandemic app, which is in accordance with previous studies on adoption rates

XSL•FO
RenderX

in 2020 in Germany [33], thus also indicating a plateau in app uptake [53]. But since research via pandemic apps can benefit from a distinctly smaller uptake rate, such as 20% to 40%, a lower app uptake is still productive. However, trustworthiness remains an important component—also when considering future pandemic apps that address more local outbreaks (eg, within hospitals as for multidrug-resistant infections) or for infections among socially vulnerable groups.

User Characteristics and Attitudes Toward App Usage

In line with the large longitudinal survey from Munzert et al [53], we found that, in the German context, the level of education, especially in terms of university degree(s), showed a significant impact on the uptake of app usage ($P<.001$) among smartphone users. In contrast, an immigration background significantly decreased the probability for app adoption ($P<.001$). Academic degree has a higher impact than immigration, but we must also consider that both social factors are also interfering. We take this as an indication that culturally formed preferences as well as linguistic aspects of information around apps can be important factors for usage. More research is definitively needed on this subject [54], also to develop culturally sensitive app information for a diverse population. This applies all the more as our findings showed a weak correlation between living in the eastern part of Germany and reluctance to use a pandemic app. Interestingly, this is in line with studies that showed significant differences in attitudes and behavior toward COVID-19 measures and policies between people living in East and West Germany [55,56]. For example, Fuest et al [55] tested the impact of pandemic information treatments on people residing in East and West Germany. They found that only West German citizens reacted significantly to the information, whereas East German citizens seemed far "less receptive to change their views based on information about economic or health aspects of the pandemic" [55]. This finding might also explain our result regarding the statistically lower rate of pandemic app uptake in East Germany. In general, differences in pandemic information and app uptake indicate that, even after 30 years of reunification, there are still experience-driven cultural and political differences toward governmental surveillance, tracing, and tracking measurements.

Regarding the factors of age and gender, other studies found that both were no or only weak predictors for pandemic app usage [20,23,26,32]. However, there are also studies indicating different tendencies for app adoption among different age groups [24,53,57]. Our results indicated no statistical correlation between personal affectedness by COVID-19 and pandemic app usage, which is contrary to previous studies that have suggested at least a weak significant impact of personal affection in terms of direct infection with COVID-19 on app adoption [20,32,58].

Two Patterns of Attitudes: Engagement Versus Privacy-Concerned Skepticism

Our study indicates a basic typology differing between users and nonusers of pandemic apps, which relies mainly on attitudinal features and less on sociodemographic factors. Type one—the data sharer—is characterized by high trust in state-funded research institutions and app providers, high willingness to share data, and seeing pandemic apps as useful for pandemic research as well as agreeing that there is a societal duty to share data to help with pandemic containment. The other type—the data-sharing skeptic—can be characterized by lower trust in state-funded app providers, decreased willingness for data sharing with research organizations, and considerably lower agreement with the view that using pandemics apps is a societal duty. These empirical findings can help to improve our understanding of who future app researchers would want to address. As problematized in other fields such as organ donation [59], technological skepticism among participants cannot sufficiently be explained by an information deficit. Hence, activities to increase the willingness to share data with research institutes might benefit from focusing on those willing to share or those who are yet undecided.

Limitations

The findings in this report are subject to at least 3 limitations. First, with regard to inclusion criteria, the participants of the survey were all residents of Germany aged 18 years or older (which is also a common ethical-legal requirement for this type of survey) and accessible via a landline or mobile phone number, so no statements can be made for people younger than 18 years or people without any telephone connection or using call blockers. Since, for example, the national pandemic app (Corona Warn App) is available for teenagers aged 16 years and older, our sample of smartphone and app users is not exactly representative of all potential app users. However, since age was not a statistically significant factor for app uptake in our survey, the question arises whether including younger populations would have had statistically significant effects on public attitudes toward pandemic app usage. Regarding people using call blockers or people without landline or cell phone numbers, we could only speculate that these populations may have a rather skeptical attitude toward sharing app data.

Second, our decision to focus our statistical analysis on smartphone users (n=778) was based on considerations that eventually non-smartphone users may not have accurate conceptions and no concrete opinions about specific applications and app data, so answers by non-smartphone users about app details could have had a rather speculative character. The characteristics of the smartphone user sample slightly differ from the German population in 3 aspects: Smartphone users are somewhat younger (–3.35 years), slightly higher educated (4.4% more with A-Level or university degree), and slightly more often personally affected by current COVID-19 disease (3.7% more are affected). Therefore, the generalizability of the present results to older people and the overall German population is limited. Nonetheless, we consider our sample more informative for app developers and governance policies than surveys based on online panels involving a convenience sampling.

Finally, due to the time limitation for telephone surveys, we opted to not provide a definition for the "use" of smartphones and pandemic apps in our questionnaire (see Multimedia Appendix 1). As no specific criterion for usage was given, the interpretation of using a smartphone or using a pandemic app was made by the respondents. Time constraints also prevented

XSL•FO
RenderX

us from asking participants about their usage of various, specific apps.

Conclusions

The rapidly expanding field of apps in mHealth is very diverse with respect to architecture, features, and purposes. Smartphones users might be confused about different types of pandemic apps [17]. Our study focused on the potential for ROA—a relatively new field with high potential to become relevant for public health research and policy making on public health. Current app development is accompanied by governance policies and ELSI (Ethical, Legal, Social Issues) research. These frameworks already consider privacy and data safety perception of the broad population as key issues.

Social Implications for Governance of App Data

Our study indicated that trust in and trustworthiness of different app providers for data storage and app distribution, self-determination of data storage and transmission, and the social attitudes toward pandemic management are also crucial for such governance. Furthermore, lay-accessible information—also considering various sociocultural groups and different levels of eHealth literacy—should be part of future frameworks. Future research, (eg, on the incentivization of app adoption and data sharing or "data donation" [53,60]) might also evaluate to what extent trust and trustworthiness can be understood as an indirect incentive and what kind of incentivization is politically and ethically justifiable.

Ethical Implications for Pandemic App Development

In order not to gamble away the high willingness to share data via an app with state-funded research institutes, the life cycle of pandemic apps and all organizational providers involved in it should be made transparent.

From an ethical point of view, public-private partnerships for app development and app operation might be reconsidered because public and private app providers are perceived very differently among smartphone users. This applies all the more when it comes to public health emergencies such as pandemics when digital solutions are rapidly recommended to fix challenges in management and containment. At least, we assume, transparency of the engaged sectors and parties can help to engage as many citizens as necessary for valid ROA deployment.

Acknowledgments

The interview study was conducted as part of the Coordination on mobile pandemic apps best practice and solution sharing (COMPASS) project. We particularly thank Johannes Huxoll from Kantar GmbH for his valuable support in pilot-testing the questionnaire and providing supplementary methodological material on the survey data. We especially thank Rüdiger Pryss and Dagmar Krefting for valuable input in planning the study and discussing preliminary results. We acknowledge Theresa Petzold for her support with the first descriptive statistical analysis and with data maps as illustrations. Ruben Sakowsky and Julia Perry deserve thanks for their thorough language editing. The COMPASS project is part of the Nationales Forschungsnetzwerk der Universitätsmedizin zu COVID-19 (NaFoUniMedCovid19) and is funded by the German Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung [BMBF]; Grant number 01KX2021). The BMBF had no role in the study design, conduct, nor writing of the article. The work is licensed under the Creative Commons Attribution License CC-BY 4.0.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire translated into English.
[PDF File (Adobe PDF File), 246 KB - mhealth_v10i1e31857_app1.pdf]

Multimedia Appendix 2
Weighting model.
[PDF File (Adobe PDF File), 173 KB - mhealth_v10i1e31857_app2.pdf]

Multimedia Appendix 3
Flow chart: recruitment and sample.
[PDF File (Adobe PDF File), 160 KB - mhealth_v10i1e31857_app3.pdf]

Multimedia Appendix 4
Attitudes among people willing to share data ("data sharers," n=653) for research via an app and people not willing to share data for research via an app ("non-data sharers," n=125) among smartphone users (n=778).
[PDF File (Adobe PDF File), 161 KB - mhealth_v10i1e31857_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5
Pandemic app use in West and East German Federal States.
[PDF File (Adobe PDF File), 419 KB - mhealth_v10i1e31857_app5.pdf]

References

1. Commission Recommendation (EU) 2020/518 of 8 April 2020 on a common Union toolbox for the use of technology and data to combat and exit from the COVID-19 crisis, in particular concerning mobile applications and the use of anonymised mobility data. EUR-Lex. 2020 Apr 08. URL: http://data.europa.eu/eli/reco/2020/518/oj/eng [accessed 2021-06-23]
2. Commission Implementing Decision (EU) 2020/1023 of 15 July 2020 amending Implementing Decision (EU) 2019/1765 as regards the cross-border exchange of data between national contact tracing and warning mobile applications with regard to combatting the COVID-19 pandemic (Text with EEA relevance). EUR-Lex. 2020 Jul 15. URL: http://data.europa.eu/eli/dec_impl/2020/1023/oj/eng [accessed 2021-06-23]
3. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet Digit Health 2020 Aug;2(8):e425-e434 [FREE Full text] [doi: 10.1016/S2589-7500(20)30137-0] [Medline: 32835200]
4. Lo B, Sim I. Ethical framework for assessing manual and digital contact tracing for COVID-19. Ann Intern Med 2021 Mar;174(3):395-400 [FREE Full text] [doi: 10.7326/M20-5834] [Medline: 33076694]
5. Ranisch R, Nijsingh N, Ballantyne A, van Bergen A, Buyx A, Friedrich O, et al. Digital contact tracing and exposure notification: ethical guidance for trustworthy pandemic management. Ethics Inf Technol 2020 Oct 21:1-10 [FREE Full text] [doi: 10.1007/s10676-020-09566-8] [Medline: 33106749]
6. Sharon T. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. Ethics Inf Technol 2020 Jul 18:1-13 [FREE Full text] [doi: 10.1007/s10676-020-09547-x] [Medline: 32837287]
7. Klingwort J, Schnell R. Critical limitations of digital epidemiology: why COVID-19 apps are useless. Survey Research Methods 2020;14(2):95-101 [FREE Full text] [doi: 10.18148/srm/2020.v14i2.7726]
8. Lucivero F, Hallowell N, Johnson S, Prainsack B, Samuel G, Sharon T. COVID-19 and contact tracing apps: ethical challenges for a social experiment on a global scale. J Bioeth Inq 2020 Dec 25;17(4):835-839 [FREE Full text] [doi: 10.1007/s11673-020-10016-9] [Medline: 32840842]
9. Cohen IG, Gostin LO, Weitzner DJ. Digital smartphone tracking for COVID-19: public health and civil liberties in tension. JAMA 2020 Jun 16;323(23):2371-2372. [doi: 10.1001/jama.2020.8570] [Medline: 32459289]
10. Proposal for a regulation of the European Parliament and of the Council on a framework for the issuance, verification and acceptance of interoperable certificates on vaccination, testing and recovery to facilitate free movement during the COVID-19 pandemic (Digital Green Certificate). EUR-Lex. 2021 Mar 17. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0130 [accessed 2021-06-26]
11. Ihre Datenspende gegen Corona. Corona-Datenspende. URL: https://corona-datenspende.de/ [accessed 2021-06-23]
12. Loosman I. 10 Rethinking Consent in mHealth: (A) Moment to Process. In: Haltaufderheide J, Hovemann J, Vollmann J, editors. Aging between Participation and Simulation. Berlin, Germany: De Gruyter; 2020:159-170.
13. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
14. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
15. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
16. Birkhoff SD, Smeltzer SC. Perceptions of smartphone user-centered mobile health tracking apps across various chronic illness populations: an integrative review. J Nurs Scholarsh 2017 Jul;49(4):371-378. [doi: 10.1111/jnu.12298] [Medline: 28605151]
17. Zimmermann BM, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. J Med Internet Res 2021 Feb 08;23(2):e25525 [FREE Full text] [doi: 10.2196/25525] [Medline: 33503000]
18. Horstmann KT, Buecker S, Krasko J, Kritzler S, Terwiel S. Who does or does not use the 'Corona-Warn-App' and why? Eur J Public Health 2021 Feb 01;31(1):49-51 [FREE Full text] [doi: 10.1093/eurpub/ckaa239] [Medline: 33340328]
19. Blom AG, Wenz A, Cornesse C, Rettig T, Fikel M, Friedel S, et al. Barriers to the large-scale adoption of a COVID-19 contact tracing app in Germany: survey study. J Med Internet Res 2021 Mar 02;23(3):e23362 [FREE Full text] [doi: 10.2196/23362] [Medline: 33577466]
20. Becker S, Degeling M, Dürmuth M, Farke FM, Schaewitz L, Schnitzler T, et al. Akzeptanz von Corona-Apps in Deutschland vor der Einführung der Corona-Warn-App. Ruhr-Universität Bochum. 2020 Jun 29. URL: https://www.mobsec.ruhr-uni-bochum.de/media/mobsec/veroeffentlichungen/2020/06/29/corona_apps_de1_preprint_de.pdf [accessed 2021-06-23]

XSL•FO
RenderX

21. Wiertz C, Banerjee A, Acar OA, Ghosh A. Predicted adoption rates of contact tracing app configurations - insights from a choice-based conjoint study with a representative sample of the UK population. SSRN Journal 2020:1 [FREE Full text] [doi: 10.2139/ssrn.3589199]
22. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in Switzerland: panel survey analysis. JMIR Public Health Surveill 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: 10.2196/25701] [Medline: 33326411]
23. Maytin L, Maytin J, Agarwal P, Krenitsky A, Krenitsky J, Epstein RS. Attitudes and perceptions toward COVID-19 digital surveillance: survey of young adults in the United States. JMIR Form Res 2021 Jan 08;5(1):e23000 [FREE Full text] [doi: 10.2196/23000] [Medline: 33347420]
24. Public perceptions of health and social care in light of COVID-19 (November 2020). The Health Foundation. 2021 Jan. URL: https://www.health.org.uk/publications/public-perceptions-of-health-and-social-care-in-light-of-covid-19-november-2020 [accessed 2021-06-23]
25. Simko L, Calo R, Roesner F, Kohno T. COVID-19 Contact Tracing and Privacytudying Opinion and Preferences. arXiv. 2020 Dec 17. URL: http://arxiv.org/abs/2005.06056 [accessed 2020-10-27]
26. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
27. Kaptchuk G, Goldstein DG, Hargittai E, Hofman J, Redmiles EM. How good is good enough for COVID19 apps? The influence of benefits, accuracy, and privacy on willingness to adopt. arXiv. 2020 May 18. URL: http://arxiv.org/abs/2005.04343 [accessed 2021-06-23]
28. Hargittai E, Redmiles EM, Vitak J, Zimmer M. Americans' willingness to adopt a COVID-19 tracking app. FM 2020 Oct 06:1. [doi: 10.5210/fm.v25i11.11095]
29. Utz C, Becker S, Schnitzler T, Farke FM, Herbert F, Schaewitz L, et al. Apps Against the Spread: Privacy Implications and User Acceptance of COVID-19-Related Smartphone Apps on Three Continents. 2021 Presented at: CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445517]
30. Lucivero F, Marelli L, Hangel N, Zimmermann BM, Prainsack B, Galasso I, et al. Normative positions towards COVID-19 contact-tracing apps: findings from a large-scale qualitative study in nine European countries. Critical Public Health 2021 Jun 02:1-14. [doi: 10.1080/09581596.2021.1925634]
31. Kaspar K. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: quantitative survey study. J Med Internet Res 2020 Aug 27;22(8):e21613 p. 10,11 [FREE Full text] [doi: 10.2196/21613] [Medline: 32759100]
32. Buder F, Dieckmann A, Manewitsch V, Dietrich H, Wiertz C, Banerjee A, et al. Adoption Rates for Contact Tracing App Configurations in Germany. Nuremberg Institute for Market Decisions. 2020 Jun. URL: https://www.nim.org/sites/default/files/medien/359/dokumente/2020_nim_report_tracing_app_adoption_fin_0.pdf [accessed 2021-12-05]
33. Corona-Warn-App: Einstellungen und Akzeptanz der Bevölkerung. Vorabergebnisse aus dem eGovernment MONITOR 2020 zur Corona-Warn-App in Deutschland inklusive Folgebefragung (Sept 3, 2020). Initiative D21 e. V. und TU München. URL: https://initiatived21.de/app/uploads/2020/09/ergebnisse-befragung-corona-warn-app-egovernment-monitor-2020-1.pdf [accessed 2020-10-27]
34. Nicholas J, Shilton K, Schueller SM, Gray EL, Kwasny MJ, Mohr DC. The role of data type and recipient in individuals' perspectives on sharing passively collected smartphone data for mental health: cross-sectional questionnaire study. JMIR Mhealth Uhealth 2019 Apr 05;7(4):e12578 [FREE Full text] [doi: 10.2196/12578] [Medline: 30950799]
35. Beierle F, Tran VT, Allemand M, Neff P, Schlee W, Probst T, et al. What data are smartphone users willing to share with researchers? J Ambient Intell Human Comput 2019 Jun 20;11(6):2277-2289. [doi: 10.1007/s12652-019-01355-6]
36. Richter G, Borzikowsky C, Lesch W, Semler SC, Bunnik EM, Buyx A, et al. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. Eur J Hum Genet 2021 Mar;29(3):495-502 [FREE Full text] [doi: 10.1038/s41431-020-00735-3] [Medline: 33005018]
37. Gesellschaft für Sozialforschung und statistische Analysen mbH. forsa. URL: https://www.forsa.de/methoden/ [accessed 2021-08-26]
38. Scherr TF, DeSousa JM, Moore CP, Hardcastle A, Wright DW. App use and usability of a barcode-based digital platform to augment COVID-19 contact tracing: postpilot survey and paradata analysis. JMIR Public Health Surveill 2021 Mar 26;7(3):e25859 [FREE Full text] [doi: 10.2196/25859] [Medline: 33630745]
39. Krutzinna J, Floridi L, editors. The Ethics of Medical Data Donation. Cham, Switzerland: Springer; 2019.
40. Mello MM, Lieou V, Goodman SN. Clinical trial participants' views of the risks and benefits of data sharing. N Engl J Med 2018 Jun 07;378(23):2202-2211 [FREE Full text] [doi: 10.1056/NEJMsa1713258] [Medline: 29874542]
41. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 contact tracing and data protection can go together. JMIR Mhealth Uhealth 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: 10.2196/19359] [Medline: 32294052]
42. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]

XSL•FO
RenderX

43. Levin-Zamir D, Bertschi I. Media health literacy, eHealth literacy, and the role of the social environment in context. Int J Environ Res Public Health 2018 Aug 03;15(8):1643 [FREE Full text] [doi: 10.3390/ijerph15081643] [Medline: 30081465]
44. El Benny M, Kabakian-Khasholian T, El-Jardali F, Bardus M. Application of the eHealth literacy model in digital health interventions: scoping review. J Med Internet Res 2021 Jun 03;23(6):e23473 [FREE Full text] [doi: 10.2196/23473] [Medline: 34081023]
45. Knitza J, Simon D, Lambrecht A, Raab C, Tascilar K, Hagen M, et al. Mobile health usage, preferences, barriers, and eHealth literacy in rheumatology: patient survey study. JMIR Mhealth Uhealth 2020 Aug 12;8(8):e19661 [FREE Full text] [doi: 10.2196/19661] [Medline: 32678796]
46. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. Journal of Media Psychology 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
47. Vaudenay S. Centralized or Decentralized? The Contact Tracing Dilemma. Infoscience. URL: https://infoscience.epfl.ch/record/277809 [accessed 2021-06-29]
48. Criddle C, Kelion L. Coronavirus contact-tracing: World split between two types of app. BBC News. 2020 May 07. URL: https://www.bbc.com/news/technology-52355028 [accessed 2021-06-23]
49. Exposure Notification API launches to support public health agencies. Google. 2020 May 20. URL: https://www.blog.google/inside-google/company-announcements/apple-google-exposure-notification-api-launches/ [accessed 2021-06-23]
50. Taylor E, Hakmeh J, Peters A, Ignatidou S. 02 The COVID-19 app: how big tech outwitted the UK government. Chatham House. 2021 Feb 16. URL: https://www.chathamhouse.org/sites/default/files/2021-02/2021-02-16-covid-19-trends-technology-hakmeh-et-al.pdf [accessed 2021-12-05]
51. Anglemyer A, Moore TH, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database Syst Rev 2020 Aug 18;8:CD013699 [FREE Full text] [doi: 10.1002/14651858.CD013699] [Medline: 33502000]
52. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. Lancet Digit Health 2020 Nov;2(11):e607-e621 [FREE Full text] [doi: 10.1016/S2589-7500(20)30184-9] [Medline: 32839755]
53. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
54. Peltonen E, Lagerspetz E, Hamberg J, Mehrotra A, Musolesi M, Nurmi P, et al. The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage. 2018 Presented at: 20th International Conference on Human-Computer Interaction with Mobile Devices and Services; September 3-6, 2018; Barcelona, Spain URL: https://dlnext.acm.org/doi/10.1145/3229434.3229474 [doi: 10.1145/3229434.3229474]
55. Fuest C, Immel L, Neumeier F, Peichl A. Does expert information affect citizens' attitudes toward corona policies? Evidence from Germany. SSRN Journal 2021:20. [doi: 10.2139/ssrn.3832479]
56. Schmelz K. Enforcement may crowd out voluntary support for COVID-19 policies, especially where trust in government is weak and in a liberal society. Proc Natl Acad Sci U S A 2021 Jan 05;118(1):1 [FREE Full text] [doi: 10.1073/pnas.2016385118] [Medline: 33443149]
57. Kawakami N, Sasaki N, Kuroda R, Tsuno K, Imamura K. The effects of downloading a government-issued COVID-19 contact tracing app on psychological distress during the pandemic among employed adults: prospective study. JMIR Ment Health 2021 Jan 12;8(1):e23699 [FREE Full text] [doi: 10.2196/23699] [Medline: 33347424]
58. Zhang B, Kreps S, McMurry N, McCain RM. Americans' perceptions of privacy and surveillance in the COVID-19 pandemic. PLoS One 2020;15(12):e0242652 [FREE Full text] [doi: 10.1371/journal.pone.0242652] [Medline: 33362218]
59. Schicktanz S, Pfaller L, Hansen S, Boos M. Attitudes towards brain death and conceptions of the body in relation to willingness or reluctance to donate: results of a student survey before and after the German transplantation scandals and legal changes. J Public Health 2017 Jan 22;25(3):249-256 [FREE Full text] [doi: 10.1007/s10389-017-0786-3]
60. Fast V, Schnurr D. Incentivizing data donations and the adoption of COVID-19 contact-tracing apps: a randomized controlled online experiment on the German Corona-Warn-App. SSRN Journal 2021:1 [FREE Full text] [doi: 10.2139/ssrn.3786245]

Abbreviations

API: application programming interface
BMBF: Bundesministerium für Bildung und Forschung
COMPASS: Coordination on mobile pandemic apps best practice and solution sharing
eHEALs: eHealth Literacy Scale
ELSI: Ethical, Legal, Social Issues
IT: information technology
mHealth: mobile health
NaFoUniMedCovid19: Forschungsnetzwerk der Universitätsmedizin zu COVID-19
ROA: research-orientated app

XSL•FO
RenderX

Edited by L Buis; submitted 07.07.21; peer-reviewed by M Lotto, C Moore; comments to author 05.08.21; revised version received 31.08.21; accepted 27.10.21; published 24.01.22.

Please cite as:
Buhr L, Schicktanz S, Nordmeyer E
Attitudes Toward Mobile Apps for Pandemic Research Among Smartphone Users in Germany: National Survey
JMIR Mhealth Uhealth 2022;10(1):e31857
URL: https://mhealth.jmir.org/2022/1/e31857
doi:10.2196/31857
PMID:35072646

©Lorina Buhr, Silke Schicktanz, Eike Nordmeyer. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 24.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

App-Based Relaxation Exercises for Patients With Chronic Neck Pain: Pragmatic Randomized Trial

Daniel Pach[1,2], MD; Susanne Blödt[1], PhD; Jiani Wang[1], MSc; Theresa Keller[3], PhD; Beatrice Bergmann[1], MD; Alizé A Rogge[1], MSc; Jürgen Barth[2], PhD; Katja Icke[1]; Stephanie Roll[1], PhD; Claudia M Witt[1,2,4], MD, MBA

[1]Institute of Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[2]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland

[3]Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[4]Center for Integrative Medicine, University of Maryland School of Medicine, Baltimore, MD, United States

Corresponding Author:
Claudia M Witt, MD, MBA
Institute for Complementary and Integrative Medicine
University Hospital Zurich and University of Zurich
Sonneggstrasse 6
Zurich, 8091
Switzerland
Phone: 41 44 2552396
Email: claudia.witt@uzh.ch

Abstract

Background: Chronic neck pain is a highly prevalent condition. Learning a relaxation technique is recommended by numerous guidelines for chronic neck pain. Smartphone apps can provide relaxation exercises; however, their effectiveness, especially in a self-care setting, is unclear.

Objective: The aim of this pragmatic randomized trial is to evaluate whether app-based relaxation exercises, including audio-based autogenic training, mindfulness meditation, or guided imagery, are more effective in reducing chronic neck pain than usual care alone.

Methods: Smartphone owners aged 18 to 65 years with chronic (>12 weeks) neck pain and the previous week's average neck pain intensity ≥4 on the Numeric Rating Scale (0=no pain to 10=worst possible pain) were randomized into either an intervention group to practice app-based relaxation exercises or a control group (usual care and app for data entry only). For both groups, the follow-up data were collected using app-based diaries and questionnaires. The primary outcome was the mean neck pain intensity during the first 3 months based on daily measurements. Secondary outcomes included neck pain based on weekly measurements, pain acceptance, neck pain–related stress, sick-leave days, pain medication intake, and adherence, which were all measured until the 6-month follow-up. For the primary analysis, analysis of covariance adjusted for baseline neck pain intensity was used.

Results: We screened 748 participants and enrolled 220 participants (mean age 38.9, SD 11.3 years; mean baseline neck pain 5.7, SD 1.3 points). The mean neck pain intensity in both groups decreased over 3 months; however, no statistically significant difference between the groups was found (intervention: 4.1 points, 95% CI 3.8-4.4; control: 3.8 points, 95% CI 3.5-4.1; group difference: 0.3 points, 95% CI −0.2 to 0.7; P=.23). In addition, no statistically significant between-group differences regarding neck pain intensity after 6 months, responder rate, pain acceptance, pain medication intake, or sick-leave days were observed. There were no serious adverse events that were considered related to the trial intervention. In week 12, only 40% (44/110) of the participants in the intervention group continued to practice the exercises with the app.

Conclusions: The study app did not effectively reduce chronic neck pain or keep the participants engaged in exercising in a self-care setting. Future studies on app-based relaxation interventions should take into account the most recent scientific findings for behavior change techniques.

Trial Registration: ClinicalTrials.gov NCT02019134; https://clinicaltrials.gov/ct2/show/NCT02019134

International Registered Report Identifier (IRRID): RR2-10.1186/1745-6215-15-490

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2022;10(1):e31482) doi:10.2196/31482

KEYWORDS

neck pain; relaxation; RCT; mHealth; smartphone app; mobile phone

Introduction

Neck pain is a global public health issue entailing a high socioeconomic burden [1,2]; moreover, it is one of the top 5 global chronic pain conditions in terms of prevalence and cause of disability [3,4]. According to the data from the European Social Survey 2014 [5], approximately 40% of all respondents reported back or neck pain. These results indicated the highest prevalence of back or neck pain in Germany (54.05%).

In most cases, neck pain is nonspecific [1]. Hence, the treatment is complex and costly. Pharmacological approaches are often used to alleviate chronic pain; however, these approaches include possible risks of tolerance, dependence, and addiction when using opioids [6,7]. Moreover, previous research showed that exercise treatment might also be beneficial in patients with neck pain [3].

Mind–body therapies focus on the interactions among the brain, mind, body, and behavior and their effects on health and disease [8]. As components of mind–body medicine, relaxation techniques have gained wide acceptance within conventional medicine [9]. The relaxation response leads to a variety of physiological benefits that may enhance pain relief through reduced sympathetic activity, decreased muscular tension, modulated pain awareness, and increased release of endogenous opioids [10,11]. Studies directly comparing the effects of self-administered versus therapist-administered interventions found similar effects on pain reduction [12]. Moreover, according to the recent *Neck Pain Guideline* of the German Society of General Practice and Family Medicine [13], learning a relaxation technique is recommended for patients with nonspecific chronic neck pain that lasts for >12 weeks. Thus, relaxation techniques alone or in addition to conventional medical care can influence the treatment and rehabilitation of chronic neck pain. However, the accessibility of cognitive and mind–body therapies for chronic low back pain and neck pain remains a major challenge [14].

Medical smartphone apps or other mobile digital health solutions can allow easy access to self-care activities [15] and support behavior changes by incorporating features such as the provision of information, tracking of activity, or providing feedback. A review [16] identified 606 mindfulness apps; however, only 3.8% (23/606) of those apps actually provided mindfulness training, and only 1 app [17] was evaluated in a randomized controlled trial (RCT). Another review [8] on apps with self-management support functions for people with persistent pain identified only 2 evidence-based apps; however, none of them were for chronic pain.

In this study, we aim to conduct a pragmatic app-based RCT to evaluate whether app-based audio relaxation exercises are more effective in reducing chronic neck pain than usual care.

Methods

Study Design

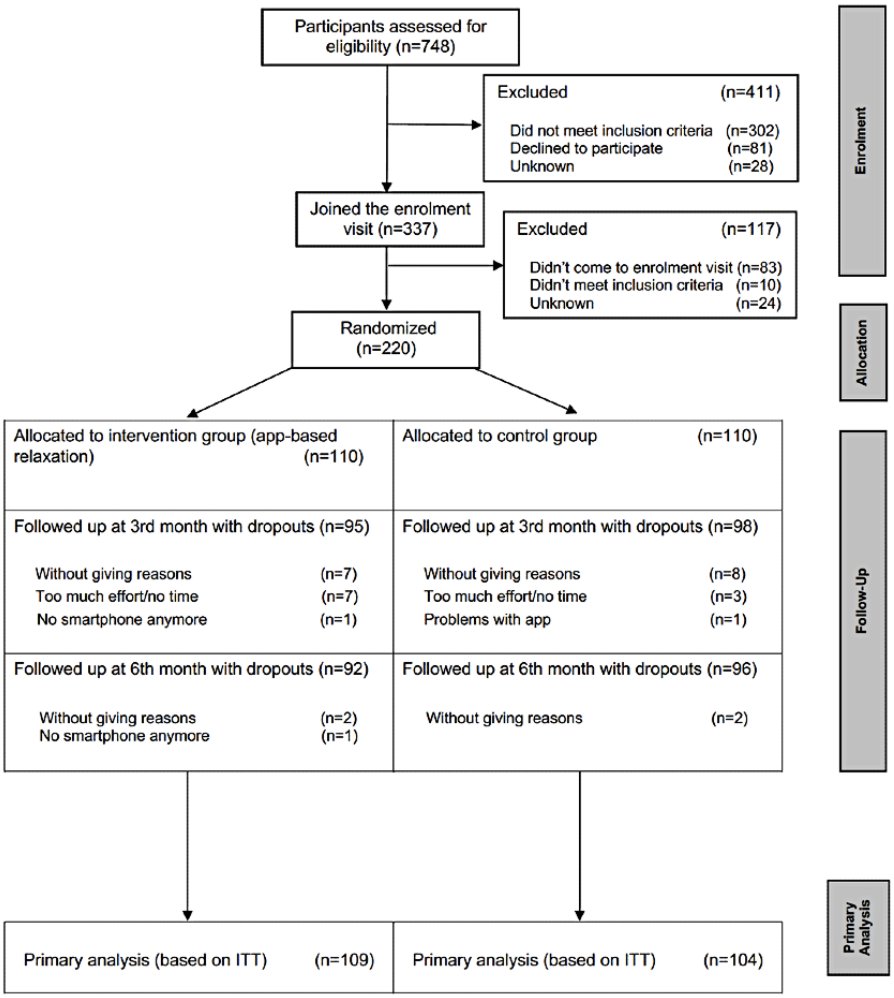
The trial design and methods have been published elsewhere [18] and have not been changed afterward. The study app remained *frozen* without any updates during the trial.

We conducted a 2-armed, randomized, parallel-group, single-center pragmatic trial to investigate the effectiveness of additional relaxation exercises delivered by a smartphone app compared with usual care alone. Participants were randomized in a 1:1 ratio to either the app-based relaxation intervention group or the control group. The trial flow is presented in Figure 1.

The intervention duration was 6 months, with the primary outcome summarizing the effect of the first 3 months.

XSL•FO
RenderX

Figure 1. Trial flow chart. ITT: intention-to-treat.



Participants and Setting

The first participant was randomized on March 31, 2014, and the final data recording was on January 11, 2017, in Berlin, Germany. Information on the study was posted with brochures and posters in universities, gyms, and general practitioners' offices. Moreover, the study was advertised in local subways from December 2014 to July 2015. Eligibility was checked by a study nurse at the study site. Eligible participants completed the paper-and-pencil baseline questionnaires. Then, the study nurse helped the participants install the app on their own smartphones and provided a randomly allocated code to activate the study app and the respective app features according to the group allocation. Participants received compensation of €20 (US $ 22.60) after participating in the study.

The inclusion criteria were as follows: aged 18-65 years, chronic neck pain within at least the past 12 weeks, average neck pain intensity ≥4 on the Numeric Rating Scale (NRS; 0=no pain to 10=worst possible pain) in the previous week, possession of a smartphone (iOS or Android), willingness to be randomized and follow the app-delivered interventions, and willingness to enter data through the study app.

Participants were excluded if their neck pain was caused by a known malignant disease, trauma, the presence of a known rheumatic disorder, a history or planned surgery of the spinal column of the lower neck in the next 6 months, known neurological symptoms (eg, radicular symptoms because of a prolapsed disk), regular intake of analgesics (more than once per week) because of additional disease, intake of centrally acting analgesics, or a history of severe acute or chronic disorders that did not allow participation in the study.

Further exclusion criteria were known alcohol or substance abuse, insufficient German language skills, current application for a pension claim, participation in another clinical trial during the 6 months before the study and parallel to the study, applying regular relaxation techniques, mindfulness meditation, or any other *mindfulness-based* therapy 6 weeks before the study or planned in the next 6 months.

Participants in both groups were allowed to continue with their usual care (medical and nonmedical); however, the regular application of any other relaxation techniques, including mindfulness meditation or mindfulness-based training, was not permitted.

The follow-up data (daily, weekly, and at the third and sixth month) were collected through the app-based questionnaires and by in-app tracking of the length of the practiced exercises. Serious adverse events were documented during the study period to evaluate safety.

XSL•FO
RenderX

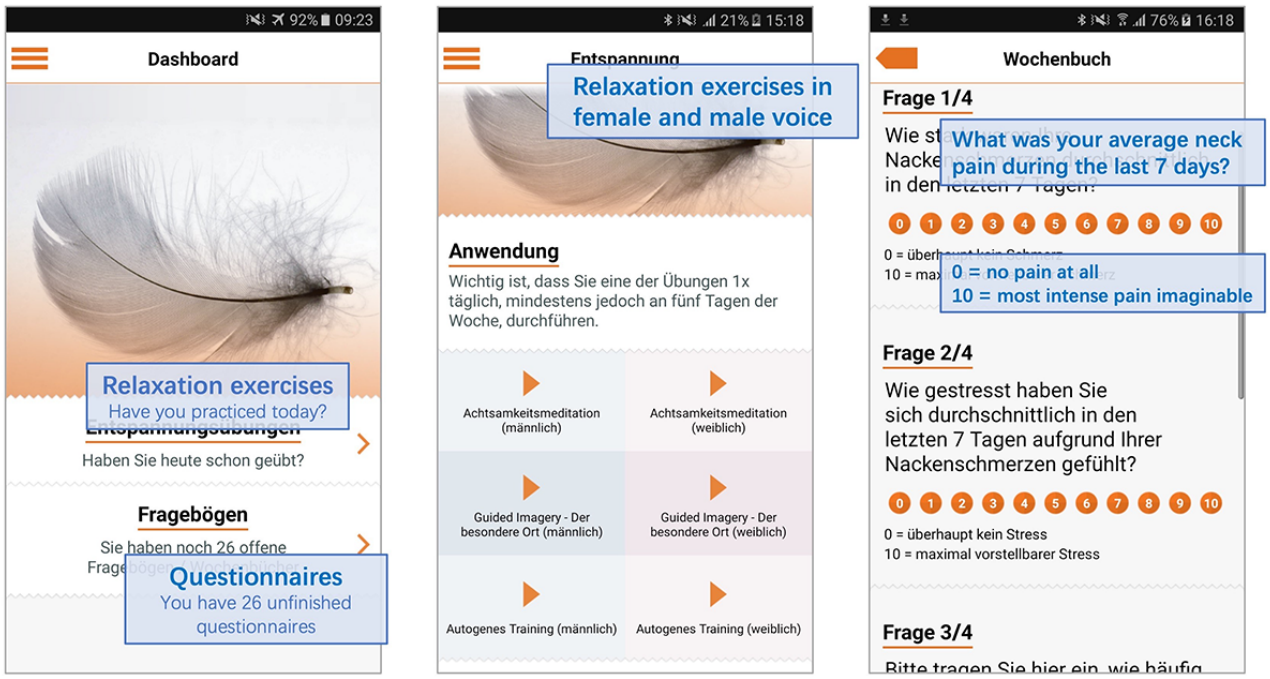
The Relaxneck App

Overview

The study app *Relaxneck* was developed by the Institute of Social Medicine, Epidemiology and Health Economics, Charité–Universitätsmedizin Berlin, Germany, together with Smart Mobile Factory, Berlin, Germany, which is an agency focused on mobile solutions [18]. The app supported iOS and Android systems and was available in the German Apple Appstore and the Google Play Store free of charge. However, the app could only be activated by entering an individual code assigned to each study participant by the study nurse.

The app supported notification features, a diary, and questionnaire options for all participants, whereas it provided audio relaxation exercises only for those in the intervention group. The app's user interface and content were available in the German language (Figure 2). The app concept was approved by the data protection officer of the Charité–Universitätsmedizin Berlin.

Figure 2. Screenshots of the study app (dashboard, relaxation exercises, and questionnaires).



App-Based Relaxation Interventions

Overview

The duration of the audios for the relaxation interventions, as well as their intensity and dosage; the use of push notifications; the diary content: and the German translation of guided imagery instructions resulted from stakeholder engagement during the planning phase of the study [18].

There were 3 types of exercises (autogenic training, mindfulness meditation, and guided imagery), with a length of 15 minutes each, that were available in 2 versions (female and male voices) in the study app for the intervention group. They were accompanied by a short instructional text (Figure 2). Relaxation exercises could be applied in different positions (sitting, walking, and lying) according to the participants' needs. It was recommended to apply a relaxation exercise daily or at least 5 days per week for 6 months.

Autogenic Training

Autogenic training is a form of self-relaxation technique that is commonly used to treat stress disorders, pain, and anxiety [19-21]. Autogenic training was developed by the German psychiatrist Johannes Schultz in 1932. It focuses on the physical sensation of the breath or heartbeat and visualizes the body as warm, heavy, or relaxed [21]. Participants learn to react to 6 verbal commands, such as "my arms are very heavy," "my heart beats regularly and calm," and "my belly is warm," to make the body feel relaxed [18].

Mindfulness Meditation

Mindfulness is a practice based on Vipassana (ie, insight) meditation, which has Buddhist roots. It is defined as "paying attention in a particular way: on purpose, in the present moment and in a nonjudgmental way" [22]. It focuses on the breath and uses it as an anchor when the mind starts to wander [18]. This concept is also used in mindfulness-based stress reduction developed by Kabat-Zinn [22-24].

Guided Imagery

In guided imagery, the mind is directed to intentionally create images to produce positive changes [25]. The audio guides the participants to visualize or conjure a place that is associated with positive feelings such as safety, security, and well-being. The guided imagery audio is accompanied by soft background music and directs visualization and imagination to a pleasant and peaceful place that has meaning for the participant to replace negative or stressful feelings [26].

XSL•FO
RenderX

Behavior Change Techniques in the App

To enhance changes in participants' behavior, behavior change techniques (BCTs) can be implemented in intervention settings [27]. As this was not a common feature in app development in 2013, we retrospectively analyzed the Relaxneck app using the BCT taxonomy (version 1) by Michie et al [27] to identify BCTs that were represented in the app, although not formally preplanned.

App for the Control Group

Participants in the control group downloaded the same app as the intervention group. All study data after baseline measurements were collected by means of app-based diaries and questionnaires. The participants were able to activate reminders for the questionnaire notifications. However, no intervention features, that is, relaxation exercises, were accessible in their version of the app. The relaxation exercises were activated after 6 months after all the survey data were collected. In addition, participants could continue using usual care, defined as all medical and nonmedical treatments, while using the app; however, relaxation techniques, mindfulness meditation, or any other mindfulness-based trainings were not permitted to be practiced during the study.

Outcome Measurements

The primary outcome measure was the mean neck pain intensity during the first 3 months of intervention based on daily measurements of pain intensity on the NRS (0=no pain to 10=worst possible pain) [18].

The secondary outcome parameters included the mean pain intensity during the first 6 months after randomization based on daily measurements, the mean pain intensity measured weekly (using NRS) as the average pain intensity of the previous 7 days over 3 and 6 months, pain acceptance (German version of Chronic Pain Acceptance Questionnaire [28]), *neck pain–related stress*, sick-leave days, and pain medication intake. Data on adherence, self-reported general changes in neck pain, suspected adverse reactions, and serious adverse events were additionally collected [18].

If a weekly survey had not been completed, the patient received an SMS text message as a reminder; if 2 consecutive weekly surveys had not been completed, the patient was contacted by telephone call; if there was no response after 2 calls, the patient received a reminder letter.

The number of participants who practiced the exercises was recorded to reflect exercise adherence over time. Practice of the exercise was defined by (1) tracking the number (and duration) of applied types of intervention with the app and (2) asking the participants weekly about the number of applied types of intervention without using the app. The complete stop of filling in any data with the study app was defined as participant dropout. Adverse events and suspected adverse reactions (only in the intervention group) were assessed after 3 and 6 months.

Sample Size

According to previous literature [29], an effect size of 0.62 has been described for mind–body therapies compared with no intervention in a group setting. We assumed a smaller effect size of 0.4 (Cohen *d*, baseline adjusted) for individual self-care relaxation exercise compared with usual care alone, as individuals might be less focused and consequently less adherent in a self-care setting [18]. To obtain a power of 80% using a 2-sided *t* test with a significance level of .05, 100 participants for each treatment group were needed (a total of 200 participants). Thus, a final sample size of 110 participants per group (220 in total), allowing a dropout rate of 9.1%, was required.

Randomization, Allocation, and Implementation

Eligible participants were randomized to either the intervention (app-based relaxation and usual care) or the control (usual care only) group using blocked randomization with variable block lengths and an allocation ratio of 1:1, that is, 110:110 participants. The randomization sequence was generated by a data manager who was not involved in the analysis of the data or the enrollment of the patients; SAS (version 9.3, SAS Inc) was used for this process. The randomization list was included in a safe Microsoft Access database to ensure that it was not accessible during the randomization process of individual participants and that the screened patients were strictly consecutively enrolled. The randomization process was conducted by the study office at the Institute of Social Medicine, Epidemiology and Health Economics. To ensure allocation concealment, first, the study team added the participants' information into the database, and then, random allocation of the participants into the intervention or control group was performed.

Statistical Analysis

For the primary analysis of the primary outcome (mean pain intensity over 3 months measured as the daily pain intensity), an analysis of covariance with a fixed factor of *treatment group*, adjusted for the baseline NRS value (fixed covariate), was performed. The analysis was based on the full analysis set (all available data without imputation of missing values, as only a small number of missing values was expected based on experiences with a previous app-based study conducted by our study team in a similar study setting [30]) based on the intention-to-treat principle with a 2-sided significance level of .05.

All the secondary analyses were explorative, and *P* values were interpreted as such. The secondary outcomes were analyzed for the full analysis set, similar to the primary analysis, depending on the scale and distribution of the outcome, that is, analysis of covariance or logistic regression, adjusted for the respective baseline value. For sensitivity analysis, the primary analysis of the primary outcome was repeated based on the per-protocol population.

Subgroup analyses were performed on the primary outcome by including an interaction term (subgroup variable by treatment) in the main model and performing separate analyses for each subgroup. Subgroups were specified with covariates in age, education (>10 years of school education or ≤10 years of school education), sex (male or female), and duration of disease. Kaplan–Meier survival analysis was conducted to investigate

XSL•FO
RenderX

whether the app features (with or without app-based intervention content) predicted the dropout of app use.

SAS version 9.4 (SAS Inc) was used for data analysis, except for the Kaplan–Meier survival analysis for adherence, which was conducted using SPSS version 22.0 (SPSS Inc).

Ethics

The study was approved by the local ethics review board at the Charité–Universitätsmedizin, Berlin (approval number Relaxneck EA 1/259/13). The study was conducted according to the common standard guidelines for clinical trials (Declaration of Helsinki and, where applicable, the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use and Good Clinical Practice revised version, Somerset West, Republic of South Africa, 1996).

All study participants provided oral and written informed consent. The trial was registered at ClinicalTrials.gov (NCT02019134), and the study protocol has been published elsewhere [18].

Results

Baseline Characteristics

Of the 748 screened participants, 220 (29.4%) were eligible for the study and gave informed consent. They were randomized either to the app-based intervention group (110/220, 50%) or to the usual care group (110/220, 50%).

The sociodemographic and clinical characteristics of the participants at baseline are presented in Table 1. The participants had a mean age of 38.9 (SD 11.3) years and an average education, with 70% (154/220) having ≥10 years of school education. Of the 220 participants, 35 (15.9%) participants had a migration background. In the previous 7 days, the average neck pain on the NRS was 5.7 (SD 1.3) points, and 26.8% (59/220) of participants had taken medication for neck pain.

Although both groups were comparable at baseline, we observed small differences regarding gender (intervention vs control: female 74/110, 67.3% vs 79/110, 71.8%), partnership status (56/110, 50.9% vs 66/110, 60%), migration background (14/110, 12.7% vs 21/110, 19.1%), duration of neck pain (mean 79.2, SD 74.8 months vs mean 86.4, SD 97.7 months), and number of sick-leave days (mean 1.7, SD 3.6 days vs mean 2.1, SD 4.5 days) after randomization.

XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of the trial groups (N=220).

Characteristics	App-based intervention (n=110)	Control (n=110)
Age (years), mean (SD)	37.9 (11)	39.8 (11.6)
Gender, n (%)		
Female	74 (67.3)	79 (71.8)
Male	36 (32.7)	31 (28.2)
BMI (kg/m²), mean (SD)	24.5 (4.6)	23.9 (4.1)
Graduation after ≥10 years of school, n (%)	79 (71.8)	75 (68.2)
Size of household, n (%)		
Single-person	32 (29.1)	34 (30.9)
2-person	44 (40)	42 (38.2)
Multiperson	34 (30.9)	34 (30.9)
Partnership, n (%)	56 (50.9)	66 (60)
Migration background[a], n (%)	14 (12.7)	21 (19.1)
Neck pain intensity in the previous 7 days (NRS[b,c]), mean (SD)	5.7 (1.4)	5.8 (1.3)
Neck pain–related stress intensity in the previous 7 days (NRS[c]), mean (SD)	5.4 (1.9)	5.3 (2.1)
Duration of neck pain (months), mean (SD)	79.2 (74.8)	86.4 (97.7)
Sick-leave days, mean (SD)	1.7 (3.6)	2.1 (4.5)
Medication intake against neck pain, n (%)	28 (25.5)	31 (28.2)
Pain acceptance, mean (SD)	73.3 (16.7)	73.6 (15.9)
Subscale pain willingness, mean (SD)	30.1 (10.1)	31.1 (8.2)
Subscale activity engagement, mean (SD)	43.2 (8.8)	42.4 (9)
Expected effectiveness of relaxation exercise, n (%)		
Recovery	1 (0.9)	5 (4.5)
Distinct improvement	54 (49.1)	61 (55.5)
Light improvement	55 (50)	44 (40)
No improvement	0 (0)	0 (0)
Ineffective	0 (0)	0 (0)
Expected effectiveness of no relaxation exercise, n (%)		
Recovery	0 (0)	1 (0.9)
Distinct improvement	3 (2.7)	6 (5.5)
Light improvement	15 (13.6)	18 (16.4)
No improvement	89 (80.9)	81 (73.6)
Ineffective	3 (2.7)	4 (3.6)

[a]Based on a study by Schenk et al [31].

[b]NRS: Numeric Rating Scale.

[c]Lower values indicate better status.

Outcomes

Less intense mean neck pain was observed in both groups during the first 3 months compared with the baseline (Table 2). However, there was no significant difference in the primary outcome of the mean neck pain intensity during the first 3 months between the intervention and control groups (group difference 0.3, 95% CI –0.2 to 0.7; *P*=.23). In addition, no significant differences in the mean neck pain intensity between the 2 groups during the second 3 months (group difference −0.1, 95% CI −0.7 to 0.4; *P*=.62) or during the entire 6 months (group difference 0.1, 95% CI –0.3 to 0.6; *P*=.62) were found.

The subgroup analysis also yielded comparable primary outcomes between participants of different genders, ages, education levels, and disease durations.

XSL•FO
RenderX

Table 2. Primary and secondary outcomes (adjusted for sex and baseline value; N=220).

Outcome	App-based intervention, mean (95% CI)	Control, mean (95% CI)	Differences intervention versus control, mean (95% CI)	*P* value
Neck pain intensity during first 3 months (NRS[a,b])	4.1 (3.8 to 4.4)	3.8 (3.5 to 4.1)	0.3 (−0.2 to 0.7)	.23
Neck pain intensity (NRS[b])				
Second 3 months	3.6 (3.2 to 4)	3.7 (3.4 to 4.1)	−0.1 (−0.7 to 0.4)	.62
First 6 months	3.9 (3.6 to 4.2)	3.8 (3.5 to 4.1)	0.1 (−0.3 to 0.6)	.62
Average neck pain during previous 7 days (NRS)				
First 3 months	4.3 (4 to 4.6)	4 (3.8 to 4.3)	0.2 (−0.2 to 0.7)	.24
Second 3 months	3.8 (3.4 to 4.1)	3.9 (3.6 to 4.3)	−0.2 (−0.7 to 0.3)	.52
First 6 months	4.1 (3.8 to 4.4)	4 (3.7 to 4.3)	0.2 (−0.3 to 0.6)	.49
Pain acceptance				
After 3rd month	75.4 (73 to 77.8)	75.8 (73.4 to 78.1)	−0.4 (−3.8 to 3)	.83
After 6th month	76.1 (73.7 to 78.4)	75.8 (73.6 to 78.1)	0.2 (−3 to 3.5)	.89
Participants with medication intake against neck pain, proportion (%)[c,d]				
During 6 months	49.5 (39.8 to 59.3)	52.4 (42.4 to 62.2)	0.97 (0.5 to 1.8)	.69
Numbers of weeks with pain medication				
First 3 months	2 (1.5 to 2.5)	2 (1.4 to 2.5)	0.01 (−0.7 to 0.8)	.98
Second 3 months	2 (1.4 to 2.6)	2 (1.5 to 2.6)	−0.03 (−0.8 to 0.8)	.93
First 6 months	3.7 (2.7 to 4.7)	3.9 (2.9 to 4.9)	−0.2 (−1.7 to 1.2)	.75
Neck pain–related stress				
First 3 months	4 (3.7 to 4.3)	3.8 (3.5 to 4.1)	0.2 (−0.2 to 0.7)	.32
Second 3 months	3.6 (3.2 to 3.9)	3.6 (3.2 to 4)	0 (−0.6 to 0.5)	.88
First 6 months	3.9 (3.6 to 4.2)	3.7 (3.4 to 4)	0.2 (−0.3 to 0.6)	.46
Responder rate, proportion (%)[c,d,e]				
After third month	29.4 (21 to 38.9)	35.6 (26.4 to 45.6)	0.75 (0.4 to 1.4)	.33
After sixth month	35.9 (26.8 to 45.7)	37.5 (28.2 to 47.5)	0.93 (0.5 to 1.7)	.80
Sick-leave days				
After third month	1.2 (0.4 to 2)	1.5 (0.7 to 2.3)	−0.3 (−1.4 to 0.9)	.66
After sixth month	1.1 (0.6 to 1.6)	1 (0.5 to 1.5)	0.1 (−0.6 to 0.8)	.81
Concomitant treatment, proportion (%)[c,d]				
After third month	40 (30.8 to 49.8)	45.5 (35.9 to 55.2)	0.82 (0.5 to 1.4)	.50
After sixth month	47.3 (37.7 to 57)	43.6 (34.2 to 53.4)	1.20 (0.7 to 2.1)	.69

[a]NRS: Numeric Rating Scale.

[b]Lower values indicate better status.

[c]Between-group differences are presented as odds ratio (95% CI) instead of mean (95% CI).

[d]Proportions are not adjusted.

[e]Either at least 50% pain reduction or at least 2.5 points on the Numeric Rating Scale compared with baseline.

Furthermore, there were no significant differences between the mean average neck pain based on weekly measurements in either group during the first 3 months (group difference 0.2, 95% CI –0.2 to 0.7; *P*=.24), second 3 months (group difference –0.2, 95% CI –0.7 to 0.3; *P*=.52), or the entire 6 months (group difference 0.2, 95% CI –0.3 to 0.6; *P*=.49).

The chance of being a responder was similar for both groups after 3 months (odds ratio 0.75, 95% CI 0.4-1.4) and after 6 months (odds ratio 0.93, 95% CI 0.5-1.7).

There were also no significant differences in pain acceptance between the groups after 3 months (group difference −0.4, 95% CI −3.8 to 3; *P*=.83) and 6 months (group difference 0.2, 95% CI −3 to 3.5; *P*=.89).

XSL•FO
RenderX

There was no significant difference between the proportions of participants who took pain medication among both groups during the whole follow-up period of 6 months (odds ratio 0.97, 95% CI 0.5-1.8; *P*=.68). The number of weeks with pain medication did not differ between the groups in the first 3 months, second 3 months, and 6 months. The number of sick-leave days and pain acceptance did not differ between the groups.

The sensitivity and subgroup analyses did not change the pattern of the results, and we found no significant difference between female and male participants in a subgroup analysis of the primary outcome.

App-Based Exercise Time and Study Dropout

The overall time spent exercising declined with time. In the first week, almost all participants (109/110, 99.1%) in the intervention group practiced the exercises with the app. However, only 40% (44/110) of the participants continued to practice the exercises (for any length) in week 12, and 30% (33/110) of the participants continued to practice the exercises (for any length) in week 26. The declining trend was similar over the study phase when comparing the number of participants who practiced relaxation exercises of any length with the number of participants who practiced relaxation exercises for at least 10 minutes per week (Figure 3).

The Kaplan–Meier survival curves in Figure 4 display the study dropouts. There was no significant difference in the curves between the 2 groups according to the log-rank test (*P*=.44).

Approximately 74.5% (82/110) of participants in the intervention group and 79.1% (87/110) of participants in the control group used the study app to answer the survey questions until the end of the study (week 26).

Figure 3. Number of participants practicing the exercises over time.

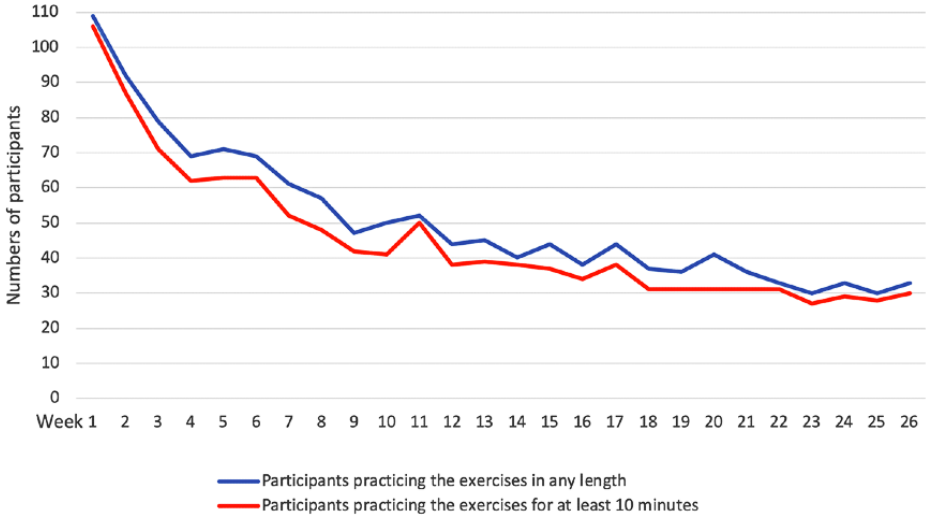
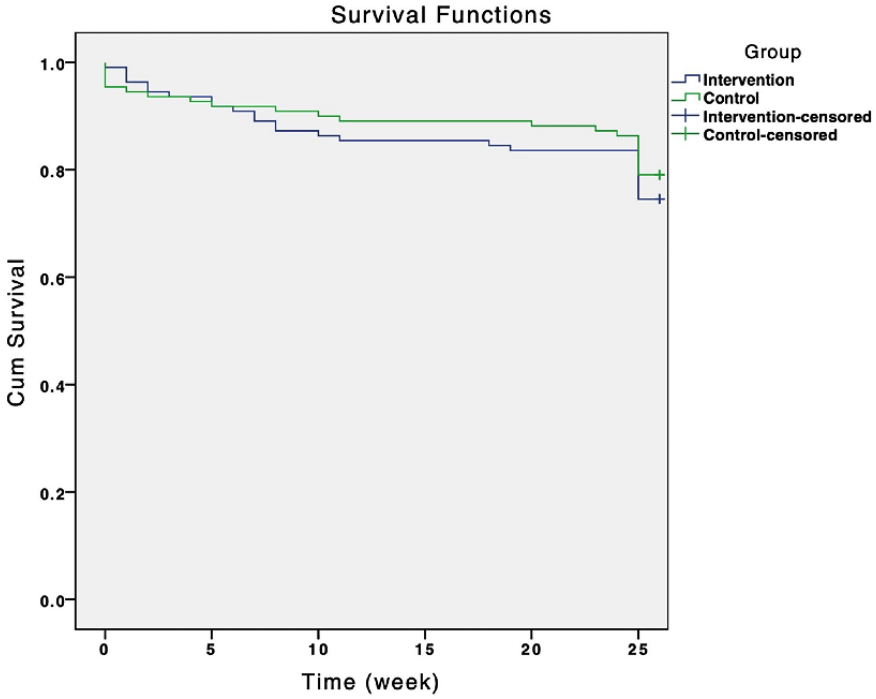


Figure 4. Probability of dropout in using the study app by group.

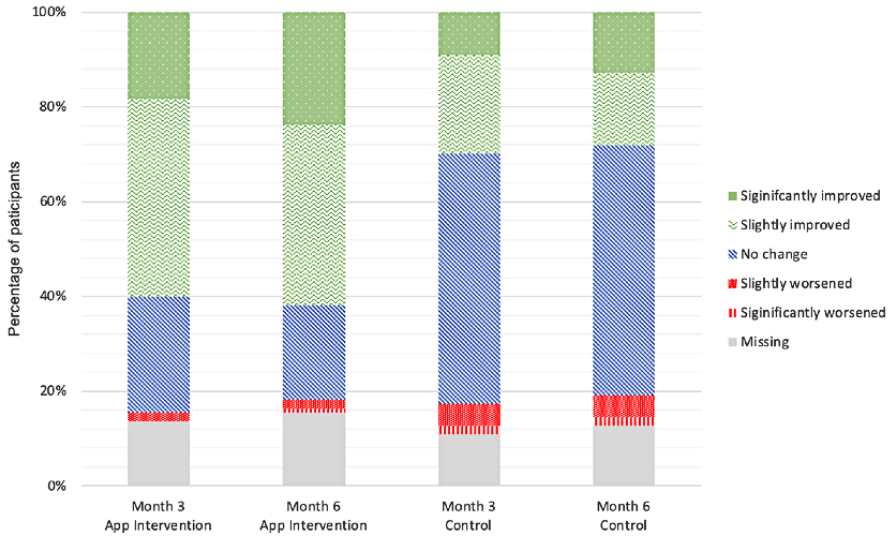


XSL•FO
RenderX

Self-perceived Neck Pain Change

Overall, 60% (66/110) of participants in the intervention group reported that they felt the neck pain improved significantly or slightly after 3 and 6 months, in contrast to approximately 30% (33/110) of participants in the control group who said the same (Figure 5).

Figure 5. Self-perceived improvement of neck pain.



BCTs in the App

Most parts of the app's user interface implementations can be characterized as *prompt and cues* BCT, such as the dashboard dialog showing the number of questionnaires remaining to be processed. Moreover, the *prompt and cues* BCT was combined with the *action planning* BCT to remind participants to fill out their weekly diaries. Participants could determine the time and date of the reminders (action planning and prompt and cues BCT).

To ensure proper performance of the relaxation exercises, all the exercises were explained by experienced clinicians in an audio recording (*instruction on how to perform the behavior* BCT). The Relaxneck app provided the full name, profession, professional title, and workplace of the audio recording instructors to ensure quality and safety for the participants (*credible source* BCT).

Safety Data

There were 5 serious adverse events recorded only in the control group, including cancer, sudden hearing loss, nerve injury and spinal tap, tonsillectomy, and an accident causing a fracture of the upper arm. None of them was considered related to the trial or the trial intervention.

Discussion

Principal Findings

In our trial, additional app-based self-relaxation techniques were not more effective than usual care alone for the reduction of chronic neck pain in a pragmatic setting. The results were consistent across all outcomes. The evaluated self-relaxation techniques were safe to use; however, they did not effectively relieve chronic neck pain during this app-based study.

There are a few possible reasons that helped to understand why the intervention did not improve pain. The study app's design was not updated during the study (developed in 2014) and did not include more elaborate BCTs, such as feedback about the correct application of the intervention and monitoring [27]. As the retrospective BCT analysis showed, only *prompt and cues* BCT was mainly used, whereas modern digital interventions or consumer apps widely apply BCTs [32,33]. In mobile health settings, personalized feedback from the app would be a promising virtual communication tool to enhance patient engagement and adherence [34]. Biofeedback and self-monitoring of changes are very important in relaxation- and mindfulness-based therapies for pain. Moreover, it must be considered that our study mainly measured self-reported outcomes. The study may have benefited from parameters such as step count as a measure of physical activity or sleep duration as a proxy for sleep [35]. At the time when the study was planned, wearables were not widely implemented, and it was more difficult to link these measures with an app because of interoperability issues. However, the type and duration of the audio recordings used as interventions were measured and used as measures of adherence. Although tracked outcomes may have added a more objective point of view, the implementation would have added a much larger complexity during the development of the app. In addition, mindfulness-based therapies are very often designed with progressive lengths or difficulties [36]. In our trial, the participants were required to practice 3 relaxation exercises of almost the same length repeatedly across the whole intervention period. This could have limited the participants' interest and the treatment effect. Finally, our app focused on audio relaxation alone instead of incorporating a whole theoretical framework such as mindfulness-based stress reduction or a comprehensive pain management strategy. Therefore, it is likely that the intervention of the study app was not powerful enough to improve chronic pain.

Adherence to the trial intervention was low compared with other app-based studies conducted by our research group [30,37]. The number of participants who performed the relaxation exercises

XSL•FO
RenderX

diminished during the course of the study. Potential explanations may again be the lack of an elaborate BCT concept or that chronic pain decreases motivation [38], especially to perform prescribed physical activities and exercises [39]. However, the number of practiced exercises of any length or >10 minutes remained similar over time. This might indicate that users who feel attached to the app-based relaxation exercise at the beginning finish the whole exercise process in most cases.

Although our study intervention was asynchronous, that is, contact with a health care provider and app intervention occurred at different time points, future mobile health studies may also include synchronous interventions in which health care providers could offer real-time interventions to the users. This approach might be helpful to improve the app and study adherence. However, this approach might also increase the complexity of the intervention and increase the costs.

In our trial, stopping the app-based intervention did not necessarily predict stopping the answering of the app-based survey questions. Only 30% (33/110) of the participants continued to practice the app-based relaxation exercises until the end of the follow-up; however, 74.5% (82/110) of participants used the app to answer survey questions until the end of the trial. Meanwhile, adherence to app use for answering survey questions was not affected by whether the app contained intervention features. The proportion of participants who used the app regularly to answer surveys until the end of the study was rather similar in both groups. A possible explanation for the good response rate in both groups could be our reminder system for the questionnaires or the paid compensation for the efforts.

Although all other outcomes did not show statistically significant group differences, most participants in the intervention group reported self-perceived improvement of neck pain, whereas most participants in the control group reported no change or worsening of neck pain. This result might be attributed to a digital placebo effect. The concept of the digital placebo effect has already been discussed in mental health studies [40]. A good example could be seen in a study involving a smartphone app that was designed to help patients self-monitor and record their symptoms of depression. Even without any direct therapeutic intervention, smartphone-based self-monitoring significantly reduced the symptoms [41]. Future studies should investigate the perceived changes in pain and the placebo-like effects of smartphone interventions.

Strengths and Limitations

Our app-based RCT was performed in a pragmatic setting. In addition, stakeholder engagement was implemented in the design of the trial and intervention [18]. Hence, the selection of the relaxation exercises and the length of the exercises were defined during stakeholder meetings to facilitate patient-centered therapy. Moreover, the study included a sufficient number of participants to answer our research question. Thus, our findings were considered generalizable in a real-life setting.

Some limitations have to be considered for this trial. The trial recruitment took rather long (32 months), possibly because of our conventional on-site recruitment strategy with paper-and-pencil baseline questionnaires. During that time, smartphone technologies, designs, and perceptions experienced numerous changes. For example, it is unclear whether the app's user interface was perceived as outdated by the participants. For future app-based studies, web-based recruitment and the incorporation of an app-based baseline survey could accelerate the overall trial process [15]. This acceleration of the trial process might also increase the relevance of the results.

Potential selection bias with an impact on the generalizability of the results might be another limitation of this study. The trial was conducted from 2014 to 2017. All study participants needed to own a smartphone. However, at that time, the number of smartphone owners in Germany (approximately 50%) was substantially lower than the current number (approximately 72%) [42]. It is unclear whether this affected the characteristics of our study population. To address a broader user base, we decided to build the study app for both the main platforms (iOS and Android).

Unfortunately, our sample size could not enable gender disaggregation. Gender might influence behavioral change, use patterns, and adherence to app use [43]. Some app-based studies have reported that gender is a strong predictor of the discontinuation of relaxation app use [37,44]. In this study, approximately 69.5% (153/220) of the participants were women. It would be interesting to discover the role of sex and gender in participants' adherence in future studies.

During the development of the app, we did not follow a preplanned BCT concept, and only basic BCTs were implemented, as shown in the post hoc review of the BCT techniques used. However, regarding behavioral change and intervention effects, a meta-analysis [45] concluded that implementing more (than one) theory is unlikely to improve intervention effectiveness. Future studies should be conducted to better understand the impact of BCTs on intervention outcomes for interventions for chronic pain.

Finally, the trial was single-blinded, as we could not blind the participants. However, it is common that participants cannot be blinded in nonpharmacological complex intervention trials and eHealth trials.

Comparison With Previous Work

Mind–body therapies are considered to be relatively safe [46]. However, only a few studies have been conducted on chronic neck pain. There were not enough trials for the Institute for Clinical and Economic Review (ICER) to summarize the effectiveness of cognitive and mind–body therapies for chronic neck pain [14]. According to a systematic review that investigated the effects of mindfulness- and relaxation-based interventions in an eHealth setting [47], only a few studies reported positive effects on pain, and no study reported positive effects on stress or mindfulness.

However, some eHealth studies have been conducted for chronic lower back pain. Heapy et al [48] reported that the efficacy of cognitive behavioral therapies (CBTs) delivered remotely using telephone and the internet for chronic back pain is not inferior to that of in-person CBTs. Kristjánsdóttir et al [49] reported that smartphone app–based interventions with personalized

XSL•FO
RenderX

feedback can reduce catastrophizing in women with chronic widespread pain. Instead of relaxation exercises alone, CBT, including emotion recognition, mindfulness exercises, and empathic communication, was highlighted in these studies. It seems that the evidence for only relaxation is rather low compared with systematic mind–body therapy or CBT for chronic pain. Therefore, future studies are required to investigate the effect of mind–body therapy on chronic neck pain within a comprehensive pain management strategy.

Conclusions

In conclusion, the evaluated study smartphone app, which included self-relaxation techniques such as autogenic training, mindfulness meditation, and guided imagery but without elaborate BCTs, was not more effective than usual care for chronic neck pain in a pragmatic trial. Further studies are needed to understand the potential of relaxation for neck pain and whether app-based mechanisms for relaxation and behavior change might be useful within a comprehensive pain management strategy for neck pain.

Acknowledgments

The authors thank Beatrice Eden and Iris Bartsch for app testing, as well as all the participants of this study. The authors thank their stakeholder team [18].

Authors' Contributions

DP, SB, and CMW conceived and designed the study. TK performed the data analysis, and JW and the other coauthors performed the data interpretation. JW and DP wrote the first draft of the paper. All the authors discussed the results, commented on the paper, and approved the final paper.

Conflicts of Interest

This was an investigator-initiated trial. The app was developed for research purposes and is not a commercial product. The authors do not have any financial stake in the success of the app. CMW received research grants from the university for digital health projects from Krebsliga Schweiz, German Cancer Aid, The German health care Innovation Fund, and Newsenselab GmbH. Board positions related to digital health for mind and body (nonpaid) are as follows: Codirector of the Digitals Society Initiative of the University Zurich and President Schweizer Fachverband Mind Body Medicine

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.2).
[PDF File (Adobe PDF File), 91 KB - mhealth_v10i1e31482_app1.pdf]

References

1. Pach D, Piper M, Lotz F, Reinhold T, Dombrowski M, Chang Y, et al. Effectiveness and cost-effectiveness of Tuina for chronic neck pain: a randomized controlled trial comparing Tuina with a no-intervention waiting list. J Altern Complement Med 2018 Mar;24(3):231-237. [doi: 10.1089/acm.2017.0209] [Medline: 29072931]
2. Hoy D, March L, Woolf A, Blyth F, Brooks P, Smith E, et al. The global burden of neck pain: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014 Jul;73(7):1309-1315. [doi: 10.1136/annrheumdis-2013-204431] [Medline: 24482302]
3. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. Mayo Clin Proc 2015 Feb;90(2):284-299. [doi: 10.1016/j.mayocp.2014.09.008] [Medline: 25659245]
4. Cohen SP, Hooten WM. Advances in the diagnosis and management of neck pain. BMJ 2017 Aug 14;358:j3221. [doi: 10.1136/bmj.j3221] [Medline: 28807894]
5. Todd A, McNamara CL, Balaj M, Huijts T, Akhter N, Thomson K, et al. The European epidemic: pain prevalence and socioeconomic inequalities in pain across 19 European countries. Eur J Pain 2019 Sep;23(8):1425-1436. [doi: 10.1002/ejp.1409] [Medline: 31038816]
6. Opioid overdose crisis. National Institute on Drug Abuse. URL: https://www.drugabuse.gov/drug-topics/opioids/opioid-overdose-crisis [accessed 2021-12-08]
7. Adriaensen H, Vissers K, Noorduin H, Meert T. Opioid tolerance and dependence: an inevitable consequence of chronic treatment? Acta Anaesthesiol Belg 2003;54(1):37-47. [Medline: 12703345]
8. Devan H, Farmery D, Peebles L, Grainger R. Evaluation of self-management support functions in apps for people with persistent pain: systematic review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e13080 [FREE Full text] [doi: 10.2196/13080] [Medline: 30747715]
9. Lin Y, Wan L, Jamison RN. Using integrative medicine in pain management: an evaluation of current evidence. Anesth Analg 2017 Dec;125(6):2081-2093. [doi: 10.1213/ANE.0000000000002579] [Medline: 29189365]
10. Sharon H, Maron-Katz A, Ben Simon E, Flusser Y, Hendler T, Tarrasch R, et al. Mindfulness meditation modulates pain through endogenous opioids. Am J Med 2016 Jul;129(7):755-758. [doi: 10.1016/j.amjmed.2016.03.002] [Medline: 27039954]

XSL•FO
RenderX

11. Zeidan F, Vago DR. Mindfulness meditation-based pain relief: a mechanistic account. Ann N Y Acad Sci 2016 Jun;1373(1):114-127. [doi: 10.1111/nyas.13153] [Medline: 27398643]
12. Palermo TM, Eccleston C, Lewandowski AS, Williams AC, Morley S. Randomized controlled trials of psychological therapies for management of chronic pain in children and adolescents: an updated meta-analytic review. Pain 2010 Mar;148(3):387-397 [FREE Full text] [doi: 10.1016/j.pain.2009.10.004] [Medline: 19910118]
13. DEGAM S1-Handlungsempfehlung Nackenschmerzen. DEGAM. 2016. URL: https://tinyurl.com/2p8432uw [accessed 2021-12-08]
14. Cherkin DC, Herman PM. Cognitive and mind-body therapies for chronic low back pain and neck pain: effectiveness and value. JAMA Intern Med 2018 Apr 01;178(4):556-557. [doi: 10.1001/jamainternmed.2018.0113] [Medline: 29507946]
15. Wang J, Rogge AA, Armour M, Smith CA, D'Adamo CR, Pischke CR, et al. International ResearchKit app for women with menstrual pain: development, access, and engagement. JMIR Mhealth Uhealth 2020 Feb 11;8(2):e14661 [FREE Full text] [doi: 10.2196/14661] [Medline: 32058976]
16. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
17. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
18. Blödt S, Pach D, Roll S, Witt CM. Effectiveness of app-based relaxation for patients with chronic low back pain (Relaxback) and chronic neck pain (Relaxneck): study protocol for two randomized pragmatic trials. Trials 2014 Dec 15;15:490 [FREE Full text] [doi: 10.1186/1745-6215-15-490] [Medline: 25511185]
19. Kim D, Rhee J, Kang SW. Reorganization of the brain and heart rhythm during autogenic meditation. Front Integr Neurosci 2014 Jan 13;7:109 [FREE Full text] [doi: 10.3389/fnint.2013.00109] [Medline: 24454283]
20. Hidderley M, Holt M. A pilot randomized trial assessing the effects of autogenic training in early stage cancer patients in relation to psychological status and immune system responses. Eur J Oncol Nurs 2004 Mar;8(1):61-65. [doi: 10.1016/j.ejon.2003.09.003] [Medline: 15003745]
21. Kanji N, White A, Ernst E. Autogenic training to reduce anxiety in nursing students: randomized controlled trial. J Adv Nurs 2006 Mar;53(6):729-735. [doi: 10.1111/j.1365-2648.2006.03779.x] [Medline: 16553681]
22. Kabat-Zinn J. Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life. New York: Hachette Books; 1994.
23. Kabat-Zinn J. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. Gen Hosp Psychiatry 1982 Apr;4(1):33-47. [doi: 10.1016/0163-8343(82)90026-3] [Medline: 7042457]
24. Kabat-Zinn J, Lipworth L, Burney R. The clinical use of mindfulness meditation for the self-regulation of chronic pain. J Behav Med 1985 Jun;8(2):163-190. [Medline: 3897551]
25. Hassed C. Mind-body therapies--use in chronic pain management. Aust Fam Physician 2013 Mar;42(3):112-117 [FREE Full text] [Medline: 23529519]
26. Relaxation techniques for health. National Center for Complementary and Integrative Health. URL: https://www.nccih.nih.gov/health/relaxation-techniques-for-health [accessed 2021-12-08]
27. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
28. Nilges P, Köster B, Schmidt CO. [Pain acceptance - concept and validation of a German version of the chronic pain acceptance questionnaire]. Schmerz 2007 Feb;21(1):57-8, 60. [doi: 10.1007/s00482-006-0508-1] [Medline: 17111168]
29. Astin JA. Mind-body therapies for the management of pain. Clin J Pain 2004;20(1):27-32. [doi: 10.1097/00002508-200401000-00006] [Medline: 14668653]
30. Blödt S, Pach D, Eisenhart-Rothe SV, Lotz F, Roll S, Icke K, et al. Effectiveness of app-based self-acupressure for women with menstrual pain compared to usual care: a randomized pragmatic trial. Am J Obstet Gynecol 2018 Dec;218(2):227.e1-227.e9 [FREE Full text] [doi: 10.1016/j.ajog.2017.11.570] [Medline: 29155036]
31. Schenk L, Bau AM, Borde T, Butler J, Lampert T, Neuhauser H, et al. [A basic set of indicators for mapping migrant status. Recommendations for epidemiological practice]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2006 Sep;49(9):853-860. [doi: 10.1007/s00103-006-0018-4] [Medline: 16927038]
32. Zens M, Woias P, Suedkamp NP, Niemeyer P. "Back on Track": a mobile app observational study using Apple's ResearchKit framework. JMIR Mhealth Uhealth 2017 Feb 28;5(2):e23 [FREE Full text] [doi: 10.2196/mhealth.6259] [Medline: 28246069]
33. Garnett C, Crane D, West R, Brown J, Michie S. Identification of behavior change techniques and engagement strategies to design a smartphone app to reduce alcohol consumption using a formal consensus method. JMIR Mhealth Uhealth 2015;3(2):e73 [FREE Full text] [doi: 10.2196/mhealth.3895] [Medline: 26123578]
34. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]

XSL•FO
RenderX

35. Carraça E, Encantado J, Battista F, Beaulieu K, Blundell J, Busetto L, et al. Effective behavior change techniques to promote physical activity in adults with overweight or obesity: a systematic review and meta-analysis. Obes Rev 2021 Jul;22 Suppl 4:e13258 [FREE Full text] [doi: 10.1111/obr.13258] [Medline: 33949778]
36. Williams JM, Russell I, Russell D. Mindfulness-based cognitive therapy: further issues in current evidence and future research. J Consult Clin Psychol 2008 Jun;76(3):524-529 [FREE Full text] [doi: 10.1037/0022-006X.76.3.524] [Medline: 18540746]
37. Mikolasek M, Witt CM, Barth J. Adherence to a mindfulness and relaxation self-care app for cancer patients: mixed-methods feasibility study. JMIR Mhealth Uhealth 2018 Dec 06;6(12):e11271 [FREE Full text] [doi: 10.2196/11271] [Medline: 30522990]
38. Schwartz N, Temkin P, Jurado S, Lim BK, Heifets BD, Polepalli JS, et al. Chronic pain. Decreased motivation during chronic pain requires long-term depression in the nucleus accumbens. Science 2014 Aug 01;345(6196):535-542 [FREE Full text] [doi: 10.1126/science.1253994] [Medline: 25082697]
39. Joelsson M, Bernhardsson S, Larsson ME. Patients with chronic pain may need extra support when prescribed physical activity in primary care: a qualitative study. Scand J Prim Health Care 2017 Mar;35(1):64-74 [FREE Full text] [doi: 10.1080/02813432.2017.1288815] [Medline: 28277047]
40. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]
41. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
42. Number of smartphone users in Germany from January 2009 to 2020. Statista. URL: https://www.statista.com/statistics/461801/number-of-smartphone-users-in-germany/ [accessed 2021-12-08]
43. Wang J, Barth J, Göttgens I, Emchi K, Pach D, Oertelt-Prigione S. An opportunity for patient-centered care: results from a secondary analysis of sex- and gender-based data in mobile health trials for chronic medical conditions. Maturitas 2020 Aug;138:1-7 [FREE Full text] [doi: 10.1016/j.maturitas.2020.05.003] [Medline: 32631583]
44. Duman-Lubberding S, van Uden-Kraan CF, Jansen F, Witte BI, van der Velden LA, Lacko M, et al. Feasibility of an eHealth application "OncoKompas" to improve personalized survivorship cancer care. Support Care Cancer 2016 May;24(5):2163-2171 [FREE Full text] [doi: 10.1007/s00520-015-3004-2] [Medline: 26563178]
45. Prestwich A, Sniehotta FF, Whittington C, Dombrowski SU, Rogers L, Michie S. Does theory influence the effectiveness of health behavior interventions? Meta-analysis. Health Psychol 2014 May;33(5):465-474. [doi: 10.1037/a0032853] [Medline: 23730717]
46. Chen L, Michalsen A. Management of chronic pain using complementary and integrative medicine. BMJ 2017 Apr 24;357:j1284. [doi: 10.1136/bmj.j1284] [Medline: 28438745]
47. Mikolasek M, Berg J, Witt CM, Barth J. Effectiveness of Mindfulness- and Relaxation-Based eHealth Interventions for Patients with Medical Conditions: a Systematic Review and Synthesis. Int J Behav Med 2018 Feb;25(1):1-16. [doi: 10.1007/s12529-017-9679-7] [Medline: 28752414]
48. Heapy AA, Higgins DM, Goulet JL, LaChappelle KM, Driscoll MA, Czlapinski RA, et al. Interactive voice response-based self-management for chronic back pain: the COPES noninferiority randomized trial. JAMA Intern Med 2017 Jun 01;177(6):765-773 [FREE Full text] [doi: 10.1001/jamainternmed.2017.0223] [Medline: 28384682]
49. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain. part 2: 11-month follow-up results of a randomized trial. J Med Internet Res 2013 Mar 28;15(3):e72 [FREE Full text] [doi: 10.2196/jmir.2442] [Medline: 23538392]

Abbreviations

BCT: behavior change technique
CBT: cognitive behavioral therapy
ICER: Institute for Clinical and Economic Review
NRS: Numeric Rating Scale
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by L Buis; submitted 23.06.21; peer-reviewed by R Jamison, MY Wu; comments to author 17.08.21; revised version received 11.09.21; accepted 12.11.21; published 07.01.22.

Please cite as:
Pach D, Blödt S, Wang J, Keller T, Bergmann B, Rogge AA, Barth J, Icke K, Roll S, Witt CM
App-Based Relaxation Exercises for Patients With Chronic Neck Pain: Pragmatic Randomized Trial
JMIR Mhealth Uhealth 2022;10(1):e31482
URL: https://mhealth.jmir.org/2022/1/e31482
doi:10.2196/31482
PMID:34994708

©Daniel Pach, Susanne Blödt, Jiani Wang, Theresa Keller, Beatrice Bergmann, Alizé A Rogge, Jürgen Barth, Katja Icke, Stephanie Roll, Claudia M Witt. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 07.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Co-design of a Smartphone App for People Living With Dementia by Applying Agile, Iterative Co-design Principles: Development and Usability Study

Sarah Fox[1], MNeuro, PhD; Laura J E Brown[2], BSc, PhD; Steven Antrobus[3], BSc, MBA; David Brough[4], PhD; Richard J Drake[2], BSc, PhD; Francine Jury[5], MSc; Iracema Leroi[1], MD, MRCPsych; Adrian R Parry-Jones[6,7], MRCP, PhD; Matthew Machin[3], BEng

[1]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland

[2]Manchester Centre for Health Psychology, Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

[3]Division of Informatics, Imaging & Data Sciences, University of Manchester, Manchester, United Kingdom

[4]Division of Neuroscience and Experimental Psychology, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

[5]University of Manchester, Manchester, United Kingdom

[6]Manchester Centre for Clinical Neurosciences, Salford Royal NHS Foundation Trust, Salford, United Kingdom

[7]Division of Cardiovascular Sciences, University of Manchester, Manchester, United Kingdom

Corresponding Author:
Matthew Machin, BEng
Division of Informatics, Imaging & Data Sciences
University of Manchester
Oxford Road
Manchester, M13 9PL
United Kingdom
Phone: 44 (0)7917 520637
Email: matthew.machin@manchester.ac.uk

Abstract

Background: The benefits of involving those with lived experience in the design and development of health technology are well recognized, and the reporting of co-design best practices has increased over the past decade. However, it is important to recognize that the methods and protocols behind patient and public involvement and co-design vary depending on the patient population accessed. This is especially important when considering individuals living with cognitive impairments, such as dementia, who are likely to have needs and experiences unique to their cognitive capabilities. We worked alongside individuals living with dementia and their care partners to co-design a mobile health app. This app aimed to address a gap in our knowledge of how cognition fluctuates over short, microlongitudinal timescales. The app requires users to interact with built-in memory tests multiple times per day, meaning that co-designing a platform that is easy to use, accessible, and appealing is particularly important. Here, we discuss our use of Agile methodology to enable those living with dementia and their care partners to be actively involved in the co-design of a mobile health app.

Objective: The aim of this study is to explore the benefits of co-design in the development of smartphone apps. Here, we share our co-design methodology and reflections on how this benefited the completed product.

Methods: Our app was developed using Agile methodology, which allowed for patient and care partner input to be incorporated iteratively throughout the design and development process. Our co-design approach comprised 3 core elements, aligned with the values of patient co-design and adapted to meaningfully involve those living with cognitive impairments: end-user representation at research and software development meetings via a patient proxy; equal decision-making power for all stakeholders based on their expertise; and continuous user consultation, user-testing, and feedback.

Results: This co-design approach resulted in multiple patient and care partner–led software alterations, which, without consultation, would not have been anticipated by the research team. This included 13 software design alterations, renaming of the product, and removal of a cognitive test deemed to be too challenging for the target demographic.

XSL•FO
RenderX

Conclusions: We found patient and care partner input to be critical throughout the development process for early identification of design and usability issues and for identifying solutions not previously considered by our research team. As issues addressed in early co-design workshops did not reoccur subsequently, we believe this process made our product more user-friendly and acceptable, and we will formally test this assumption through future pilot-testing.

(JMIR Mhealth Uhealth 2022;10(1):e24483) doi:10.2196/24483

KEYWORDS

agile; dementia; co-design; cognition; mHealth; patient public involvement; software development; mobile phone

Introduction

Background

In January 2019, the National Health Service published its long-term plan, setting out key ambitions for the next 10 years. One of the most ambitious targets of this plan was in the field of digital technology, with a vision toward increasing care at home using remote monitoring and digital tools [1]. This move will likely be expedited by the need for social distancing brought about by the COVID-19 pandemic. Therefore, with the impetus and growing necessity for distance health care, it is important to consider how this new type of service will meet the needs of patients. A way to ensure that new technologies are usable, acceptable, and tailored toward the patients they aim to support is to ensure that the patients themselves are central to the design and development process. Co-design offers a way to ensure that new technologies and interventions are tailored to patient needs [2]. Indeed, there is growing support for the benefits of co-design in health care [3] but less evidence as to how these approaches can be tailored to the needs of diverse patient populations [4].

Within the context of software development, co-design can be defined as a process that draws on the shared creativity of software developers and people not trained in software working together [5]. To this end, special attention is given to involving end users and ensuring that their input as experts through experience is central to the design process and that their specific needs are understood and met [5-7]. This is in line with existing literature suggesting that integrating patient voice with software development is achievable and can provide valuable feedback to improve the intuitive design and usability of software outcomes [8,9].

In dementia research, patient and public involvement (PPI) and co-design is still a developing field [10], although it has been suggested to confer benefits to research outcomes, researchers, and members of the public who play a part in the process [11,12]. The capacities, needs, and preferences of those living with cognitive impairments can be diverse [13]. Therefore, standard co-design and PPI methodologies often need to be adapted to suit this population. This may be particularly important in the field of digital technology and software development, as studies suggest that there is utility and an appetite for assistive technology for older people and those living with cognitive impairments. However, despite there being motivation for older people to use digital technologies, barriers exist around usability and lack of experience [14-17].

An area in which digital technology can help with the care and management of dementia is through the monitoring of cognitive change and variability, which is an issue that is considered important for this population [18,19]. For instance, many people with dementia experience worsening cognitive and neuropsychiatric symptoms in later periods of the day, a phenomenon known as *sundowning* [20,21]. Current practice bases the diagnosis of dementia on a combination of clinical history, biomarker detection, and examination, of which cognitive testing is a key part [22]. However, conventional cognitive assessments cannot detect short-term fluctuations in cognition that might be relevant to understanding or managing individuals' cognitive, functional, and behavioral symptoms.

Recent developments in computerized cognitive testing have made it possible to measure microlongitudinal patterns of cognitive function [23]. However, although these tools have been tested in cognitively healthy older adults [14], they have not yet been used in populations with cognitive impairment. Furthermore, these tasks were designed for use on large, touchscreen tablet devices and have not yet been adapted for use on smaller, more mobile devices, such as smartphones, which are used by an increasing number of older people [23]. Therefore, there is an impetus to adapt such tasks for use on smartphone devices and meet the needs of those living with clinical conditions that affect their cognitive abilities [24].

Despite the diverse and divergent lived experiences of those living with dementia, software apps are rarely designed with this patient population in mind [15]. It is even rarer to find software codeveloped alongside those living with dementia [4]. This can result in poorer quality technology that can be difficult to use for those living with dementia [25]. However, research indicates that those living with dementia have an interest in assistive technology and are capable of using touchscreen technology [17,26]. Therefore, we approached the adaptation of microlongitudinal computerized cognitive tests to the needs of people living with cognitive impairments through an iterative Agile process with patient co-design at its center.

Agile software development focuses on collaboration with users and rapid software deployment [27]. Scaling tests to a mobile device requires regular input and development iterations from end users, with an understanding that direct translation between devices may be unsuitable. The Agile methodology is best suited to such projects where requirements may not be clearly defined at the outset and emerge over time [28]. This is especially relevant in this case, where iterative co-design workshops spaced throughout the development process meant that the final product was not clearly defined early in the process and, instead,

XSL•FO
RenderX

emerged based on consultations with experts through experience via regular workshops.

Objectives

In this paper, we describe how we modified co-design approaches to involve members of the public living with dementia and their care partners in the production of a smartphone app. Although we worked specifically with people with dementia, the principles could be applied to other patient groups who do not find it easy to engage with standard co-design approaches. We also explain the benefits of the Scrum development methodology as a way of integrating user feedback into the design and development process.

Methods

PPI and Co-design: Theoretical Framework

Overview

Public involvement in research is defined by the National Institute for Health Research's INVOLVE as research being conducted *with* or *by* members of the public rather than *to*, *about*, or *for* them. Tambuyzer et al [29] also recognize that, given the heterogeneity of research protocols and patient populations, involvement is not a *one-size-fits-all* concept and is better defined by values rather than protocols. These values include participation in decision-making and giving contributors some control and responsibility over research outcomes, active involvement that goes beyond consultation or receiving information, involvement in a range of activities, being recognized as experts by experience, and collaboration with professionals.

Working alongside individuals living with cognitive impairments necessitates a tailored approach to involvement and co-design. Therefore, it is necessary to balance facilitating meaningful involvement alongside being mindful of individuals' capacity, capability, and preferences.

Therefore, we approached the challenge of co-design alongside individuals with cognitive impairments by adopting the following three methodological steps: (1) end-user representation at research and software development meetings via a patient proxy; (2) equal decision-making power for all stakeholders based on their expertise; and (3) continual user consultation, user-testing, and feedback.

Step 1: End-user Representation

On the basis of the combination of a short timescale for app development, limitations in the availability of clinical advisors, and a desire to reduce unnecessary burden on contributors living with dementia and their care partners, we chose to represent the patient or public voice at research group meetings via a proxy. Our proxy was a PPI officer who worked alongside our research group. They were responsible for developing and facilitating co-design workshops and representing end users at research group meetings. This ensured that the patient voice was represented in all important decisions and was given equal weight as the voice of other research team members.

Step 2: Equality of Expertise

Input from those with lived experience of cognitive impairments was integral to the development of this app. Therefore, those involved were encouraged to input into all the elements of the design process. To this end, input from those with lived experience led to 13 design alterations across the life of the project (listed in the following sections). Feedback from co-design workshops also led to the removal of 1 cognitive test, which was deemed too challenging for those living with dementia, and rebranding of the app.

Step 3: Continued Input

Following the development of an initial prototype app, which was designed to act as a scaffolding example app for use in the first co-design workshop, all subsequent software development sprints were based on end-user feedback. This ensured that any emerging design or software features were reviewed and modified by end users before being added to the following sprint. The extent of the end-user modifications adopted in the creation of this app can be visualized by comparing Figure 1 (the research team's prototype app) with Figure 2 (alterations made following our first co-design workshop) and Figure 3 (the final product based on feedback from 4 co-design workshops).

XSL•FO
RenderX

Figure 1. (A) Simple test of cognitive processing speed and (B) a more cognitively demanding tests of working memory developed by the software development team before patient and public involvement input. PPI: patient and public involvement.

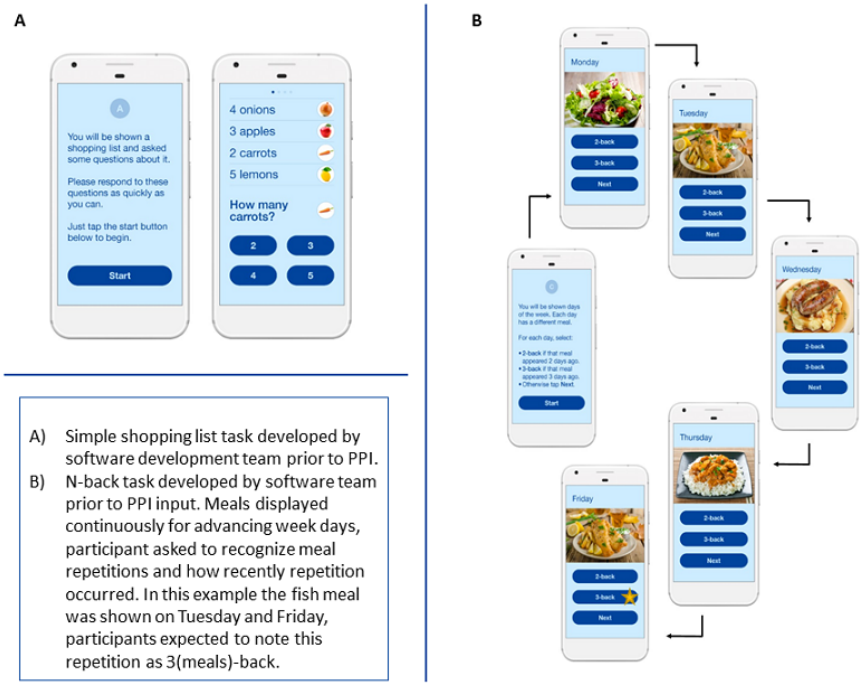
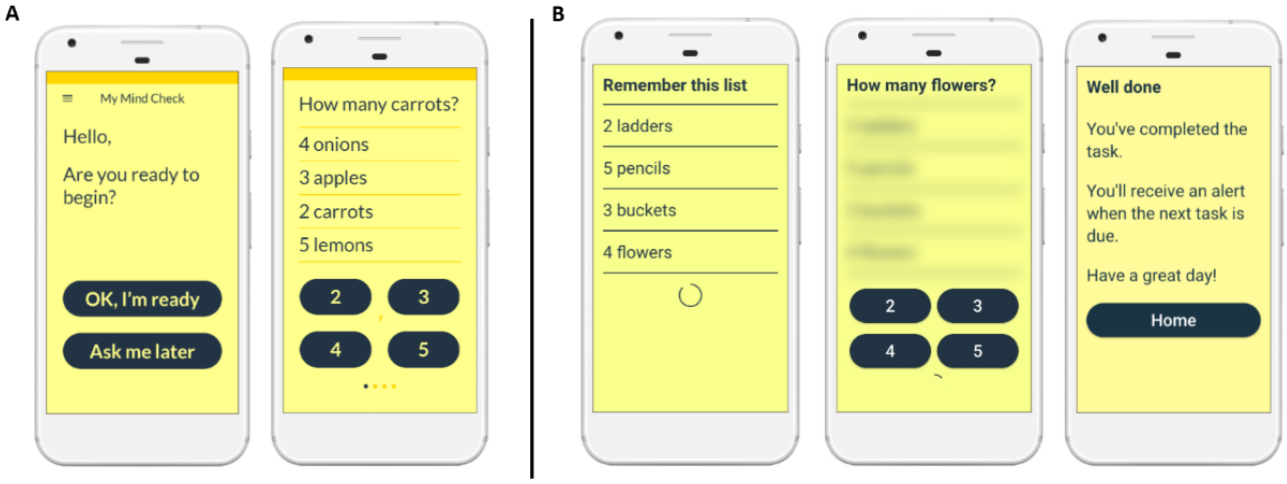


Figure 2. (A) Redesigned shopping list task and (B) new shopping list+ task following first patient and public involvement workshop.



XSL•FO
RenderX

Figure 3. Final software alterations following second patient and public involvement workshop showing flow through the app.

PPI, Co-design Process, and Methodology

Overview

A total of 4 co-design workshops were run collaboratively with community dementia support groups and were tailored to those living with cognitive impairments. Workshops were planned around familiar venues and, in some cases, to coincide with existing support group meetings. The materials used were dementia-friendly [30] and in line with INVOLVE recommendations [31]; the budget was ring-fenced to cover attendee travel, attendance fees, and refreshments.

Participants of workshops 1 and 2 comprised a mix of individuals living with a dementia diagnosis and current and past care providers of people living with dementia (workshop 1: 5/7, 71% with dementia and 2/7, 29% current or past carers; workshop 2: 3/6, 50% with dementia and 3/6, 50% current or past carers). Participants were recruited from 2 local dementia support groups following informal visits and presentations from the members of the research team. Approximately 30% (3/10) of the participants (1/3, 33% living with dementia, and 2/3, 67% current or past carers) attended both workshops 1 and 2.

Workshops 1 and 2 adopted a similar format: each workshop lasted approximately 2 hours and included (1) lunch and informal ice-breaker conversations; (2) a short, accessible project discussion and feedback; (3) introduction and testing of a visual working prototype; and (4) the collection of informal one-to-one and group feedback on the prototype. Workshop 1 also included an activity in which participants were encouraged to discuss their views on and responses to candidate words and phrases for the app's name. This was achieved via a discussion of flashcards containing keywords associated with the cognitive testing app (eg, cognition, test, training, research, brain, *e*, memory, noggin, and mobile) alongside our prototype name *Health-e-Mind*. This discussion generated the name *MyMindCheck*, which was considered meaningful and acceptable to workshop attendees. This name was later presented to the participants of workshop 2, 7 months after workshop 1, and was received positively.

These workshops were designed to involve patients in the co-design of the *MyMindCheck* app rather than being structured research or focus groups. Therefore, feedback from participants was not treated as research data. Consistent with common involvement practice [32], participant feedback was collected as written field notes by 2 workshop facilitators (direct quotes were not included); these notes were collated, and key points were identified and fed back to the research team.

Workshops 3 and 4 took place 2 months after workshop 2 and spanned a week-long period of user-testing. Participants were recruited by the research team from a local community dementia support group with a focus on technology; these were older individuals with current or past experience of supporting someone living with dementia (n=4 current or past carers). This group was targeted as we expected individuals attending a technology-focused group to be inclined to take part in our week-long user-testing phase.

Potential participants from this group were approached during one of the group's regular meetings, and the project was introduced, and the app was demonstrated. From this meeting, 4 individuals consented to the 7-day testing period, whereas 4 declined, citing time commitments as a barrier to participation. We returned to the group the following week to distribute phones preloaded with the *MyMindCheck* app, a short instruction manual, and optional paper diaries. The paper diaries were used as an aide-memoir for participants to record their day-to-day experiences using the app.

XSL•FO
RenderX

Workshop 4 took place after the week-long user-testing period and comprised a short informal discussion regarding participants' experiences of using the software. Participants were asked to comment not only on their own experience with the software but also on its suitability for someone living with a dementia diagnosis. Paper diaries were referred to during this discussion as a memory prompt; data from these diaries were not stored or analyzed further outside this workshop.

Feedback from each workshop was reviewed and discussed by the project team shortly after each workshop. This resulted in an agreed set of changes for the subsequent shippable products. As with any feedback of this nature, the project team prioritized changes based on both the effort required to implement and the likely impact on the end user.

The Software Development Process

The *MyMindCheck* app was developed using a co-design approach [5-7] involving three key groups of stakeholders: the research team (including clinical input), the target user group (people living with dementia and individuals with direct experience of caring for those with dementia), and the software development team. The software team adopted the Scrum framework for development [33]. Scrum is a modern, Agile software development methodology that fits well with the co-design approach used to develop this app. It focuses on the regular delivery of working software (*shippable products*) to users and depends on user feedback throughout the software development.

Scrum uses sprints, which are timeboxed development efforts, usually 1 to 4 weeks in duration [33]. The software team used 3-week sprints for this project, as this presented a suitable balance between the need to be able to respond flexibly to changing requirements and the delivery of sufficient functionality within each sprint. Sprint planning sessions were attended by members of the research and software development team, including a PPI specialist (research team member) who facilitated public workshops and acted as a customer proxy during these planning sessions. The customer proxy acted as the Scrum product owner in this instance and was responsible for being the voice of the customer. In Scrum, the product owner is responsible for defining and prioritizing the requirements for the product, which, in this case, was the *MyMindCheck* app [34]. The PPI specialist was chosen as the product owner as they worked closely with the PPI participants to capture the requirements for the app.

Visual working prototypes were used during the initial PPI workshop events. These prototypes allowed users to see interactive screens that portrayed key design elements and the flow through the app and were produced by a user experience designer embedded in the software team. This enabled early testing with the research team and target user group without requiring significant investment in software development.

The initial prototypes were based on a validated computerized cognitive task, in which participants were presented with a short list of grocery items (eg, carrots) and a quantity for each [23]. Participants were then asked to report the quantity of a given item (Figure 1). This task had previously been validated as a measure of cognitive processing speed in community-living older people [23]. An additional task was also included to place a higher demand on memory, which was based on clinical input. This task was based on an N-back test of working memory [35] and presented participants with a meal for each day of the week. When a participant was shown a meal that had previously been seen, they needed to recall how many days *back* the meal was first seen (Figure 1).

Before the first PPI workshop, research team members, who had experience in working with people with cognitive impairment, reviewed the prototypes. They suggested modifications to simplify these tasks, making them more visually appealing and quicker to navigate to maintain user engagement. These modifications included the following:

1. The addition of images to both tasks
2. Start screens containing instructions on how to complete the tasks
3. Simplification of the second, harder, N-back task by introducing images of meals and the use of the days of the week, as these were familiar items (Figure 1)
4. The team also generated the provisional name
   Health-e-Mind
   for the prototype app

These prototypes were then presented for review to the target user group during co-design workshops arranged by the team's public involvement lead. On the basis of initial feedback on the visual working prototypes, the software team began the development of version 1 of the app. This process continued in an iterative manner for each of the key components of the app.

Learning From the Approach

Throughout the project, the research and software development teams met on an approximately monthly basis to review the approach being taken. This enabled improvements to be made to the process within the project and resulted in some key learning points that could be applied to future projects.

Results

Initial Prototype and Feedback From Workshop 1

The initial prototype comprised a shopping list task and an N-back task (Figure 1).

Much of the participant feedback collected during workshop 1 correlated poorly with the research team's prior assumptions. Some of the key feedback from workshop 1 and the actions taken to address this feedback are presented in Table 1.

XSL•FO
RenderX

Table 1. Feedback and resulting software modifications from workshop 1.

Item and feedback	Software modification
Appearance	
• Testers found pictures to be distracting. Specifically, it was noted that some pictures were confusing (ie, onion and apple looked similar) and that the images shifted focus away from reading written information, making it harder to follow instructions.	From this feedback, the software development team chose to remove images from both tasks.
• Participants reported that the color scheme (dark blue text on a light blue background) might be inappropriate for those with reading or perceptual difficulties. Black writing on a yellow background was suggested to be optimal for improving reading speed and for assisting people with reading difficulties.	The display was altered to black text on a yellow background.
Instructions	
• Testers noted that detailed introductory text explaining the task was not necessary for the simple shopping list task. Indeed, several participants stated that they skipped reading the introductory message and were still able to perform the task.	The development team removed the introduction text from this task, replacing it with a simple “Are you ready to start <yes>, <no>” structure.
• It was noted that the screen flow used in the shopping list task left some participants confused. Specifically, several participants felt that displaying the shopping list followed by a probe question was less logical (harder to follow) than displaying the probe question first followed by the shopping list.	Text flow was altered in line with workshop preferences in the next design iteration (Figure 2).
• The shopping list task relied on measures of task completion time as a proxy for cognition. Therefore, instructions for this task asked participants to complete the task “As quickly as possible.” Workshop participants noted that although they read this instruction, they did not feel a sense of urgency while completing the task, suggesting that they had not remembered it.	To encourage users to complete the shopping list task as quickly as possible, the development team added a circular bar countdown timer to the bottom of the task screen.
N-back task	
• Participants felt that the written explanation for the second, harder (N-back) task was insufficient, and, even after a verbal explanation and demonstration, many were still uncomfortable interacting with this task.	It was decided that the N-back task was too complicated and not fit for purpose. Therefore, the development team removed this task and replaced it with a more memory-intensive variant of the shopping list task, subsequently referred to as shopping list+, in which the shopping list was removed from the screen before and during each probe question (Figure 2).
Feedback	
• Participants were asked whether they would appreciate feedback on their performance on these tasks. Opinions were mixed, with some participants wanting graphed data, or indications of low and high performance, whereas others felt that feedback on poor performance might reduce their motivation to complete future tasks.	It was decided that a generic positive feedback message would be added to the tasks, that is, “Great job, well done.”
Name	

• Participants did not like the name the research team chose for the app—Health-e-Mind. Most were unaware that the e stood for electronic, and 1 individual mentioned that it made him think of drug use. Group feedback on flashcard word association included the following:
  • Brain was seen to be too biological, whereas mind was preferred as this sounded more holistic and accessible.
  • Although some participants were comfortable with the words test and memory, others suggested that these terms may be off-putting and could cause anxiety. It was suggested that the word test could be replaced by check as this sounded less daunting and clinical.
  • Participants also liked the addition of the word my to the name, personalizing the app.
	From this feedback, the team chose to change the name to *MyMindCheck*.

XSL•FO
RenderX

Overall, workshop participants seemed positively disposed to the purpose of the app and said that assuming certain alterations were made, they would be willing to interact with such a program on a subdaily basis.

Second Prototype and Feedback From Workshop 2

Building on feedback from workshop 1, the software team undertook a second development sprint, updating the original prototype to incorporate feedback from workshop 1, including removal of N-back task and replacement with shopping list+ task (Figure 2).

Feedback from this workshop and actions taken are listed below in Table 2.

Table 2. Feedback and resulting software modifications from workshop 2.

Item and feedback	Software modification
Instructions	
For the new shopping list+ task, a number of participants noted that until they reached the screen containing the question and multiple-choice answers, they did not realize that they had to remember both the objects listed and the associated number of items.	This was addressed by altering the prompt used on the first screen of this task to read "Remember how many of each item."
Appearance	
For the shopping list+ task, several participants were unable to read the entire list of 4 items displayed on the first screen before it timed out and moved on to the probe question.	To address this, the team increased the display duration of the first screen to give users more time to read the instruction and object list. They also reduced the list length from 4 items to 3.
This version included a countdown timer on both tasks, specifically, a circular bar countdown timer. Although most testers said that they did not notice this timer, they did note that they had been trying to respond quickly. However, 1 tester did say that she noticed the timer and felt stressed about completing the task in time.	The countdown timer remains in the app as a visual cue to complete in a timely manner. However, the timer was altered from a model which showed a finite time counting down to a timer that did not count down to a finite point. It was hoped that this maintained a sense of urgency but would mitigate stress caused by a finite countdown.
Feedback	
This version of the app included a generic positive feedback message after each task that was not linked to performance, that, "Well done." This was included to avoid user discouragement because of low scores. However, participants did not appreciate being given positive feedback when they were aware that they had performed badly.	Feedback was altered to maintain a positive tone while also remaining performance neutral: "Task complete! You have finished the task. See you at the next alarm."

In line with feedback from the first workshop, no major objections were raised to the usability and acceptability of the *MyMindCheck* app. Indeed, 1 attendee who stated at the beginning of the workshop that she did not use mobile phones was particularly fast to pick up both tasks and noted at the end of the workshop that she had enjoyed testing the app.

Final Prototype and Feedback From Workshops 3 and 4

Workshops 3 and 4 aimed to test the software alterations implemented as a result of workshop 2 and to trial new functionality, including prompts and alarms. Feedback from these workshops and the actions taken to address this feedback are listed as follows:

Although most participants complied with the assigned in-app tests, the most common reasons for noncompliance were

1. The alarm was not loud or long enough.
2. Fear of breaking the phone if they took it out of the house.
3. Fatigue at being asked to complete tasks 4 times a day.

To address these concerns, the software team implemented the following modifications:

1. Increased the alarm volume and duration
2. Decided to provide phone cases when using a study phone to reduce fear of dropping or damaging phones
3. Decided to implement further PPI regarding prompt number and frequency before further implementation or testing

One tester also noted in her diary that, for the first 2 days of testing, the phone did not register her responses, and therefore was timing out on the tests. Similar issues had surfaced with other testers to a lesser extent in some of the preceding workshops. On the basis of this feedback and previous observations, it was noted that some participants were holding the response buttons rather than tapping them, perhaps reflecting a level of unfamiliarity with mobile technology among this group. Therefore, to address this issue the software was modified to identify both on-press and on-hold events as valid answers.

Testers were confident using both tasks, although they noted that they found the shopping list+ task harder and that it required more concentration. Testers with experience caring for someone with dementia stated that they believed that these tasks could be completed by someone living with dementia, assuming they had support from a care partner.

Discussion

Principal Findings

We used a co-design approach to develop the *MyMindCheck* app involving three key groups of stakeholders: the research team (including clinical input), the target user group (people living with dementia or individuals with direct experience of

XSL•FO
RenderX

caring for others with dementia), and the software development team. As patient involvement and co-design in dementia research is in its relative infancy across Europe [10], this study will make an important contribution toward a model of best practice for related research and provide an exemplar for others wishing to adopt and modify this approach. Our report conforms to the Guidance for Reporting Involvement of Patients and the Public-2 international reporting guidelines for PPI [36]. Therefore, findings from this study will be comparable and address concerns raised by some researchers regarding the lack of consistent reporting in co-design research [37], especially in regard to those living with dementia [4].

By adopting Scrum in this context, we were able to realize the benefits of an iterative co-design approach, with the software evolving throughout each of the 4 workshops. In addition, our use of prototype designs in the first 2 workshops provided the team with a low-cost opportunity to receive feedback and evaluate the idea before commencing software development.

Participants in the workshops gave positive feedback about the experience, showed strong engagement during the sessions, and provided constructive comments on the app. Notably, points raised in early workshops did not resurface in the week-long test undertaken by a different participant group at a later stage. This could be because users in the week-long test focused on different aspects of the app. However, it could also suggest that our approach was effective in addressing design issues at an early stage of development.

Although the co-design methodology enabled the team to iteratively develop the app, we still had to overcome several challenges. For instance, it was agreed early on that embedding of an end user (in this case, someone living with a dementia diagnosis into the Scrum team), as per co-design best practice [38], would not be feasible because of the burden that regular meetings could place on those living with a dementia diagnosis and their care partners, as well as the power differentials and communication difficulties associated with involving lay members in technical discussions [39]. Instead, we took the pragmatic approach of running workshops throughout the project to garner regular feedback from the user group and provide end-user representation through a proxy (in our study, the proxy was a public engagement officer who developed and facilitated all co-design workshops). Ideally, the project would have benefited from more regular contact with the end-user group. However, given the vulnerable nature of this group, our approach seemed to be an appropriate compromise, given that this was a fast turnaround, intensive development project.

There were some logistical challenges in running the Agile development using a co-design process. Specifically, recruiting participants for workshops required multiple interactions with community groups to garner interest in the project and plan suitable times and venues for workshops. Therefore, it was necessary to set the date for each workshop several weeks in advance. However, software development does not always run according to the plan, as it is not possible to estimate development tasks with a high degree of accuracy. Therefore, there is a risk that a date could be set for a workshop only for the software not to be ready in time. We mitigated this risk by setting a date for each workshop, which allowed a suitable leeway for any unexpected delays. We also worked with an experienced and established Scrum team, meaning that the estimates could usually be provided with a reasonable level of confidence.

Conclusions

Given the need for health research, particularly the development of health technology, to be approached in a patient-centered manner [37], we developed a methodology that combines Agile software development with integrated patient co-design. This approach facilitated meaningful user involvement in a manner that was easily manageable by our project team, who were working on a short timescale with budget constraints, a challenge experienced by many developers [40].

We also highlighted several instances where input provided by people with lived experience of dementia helped our team to identify and address usability issues early in the development process, speeding up delivery and reducing software development waste. Our experience evidences how co-design can benefit the software development process and be sustainably tailored to the needs of diverse patient populations [4].

The next step for the *MyMindCheck* app is to undertake a large pilot trial and adaptations to apply to other health conditions with fluctuating cognitive states. Patient groups will continue to be involved throughout this future work to ensure that the developed software is fit for its purpose.

Acknowledgments

This work was funded by the Medical Research Council Momentum Award: Institutional pump-priming award for dementia research (MC_PC_16033). The authors would like to thank the Alzheimer's Society and their patient network for support in the setup and design of workshops and all the participants who contributed to this project, including special thanks to the Peter Quinn Friendship Group, Together Dementia Support, and the Humphrey Booth Resource center.

Conflicts of Interest

None declared.

References

1. The NHS long term plan. NHS. URL: https://www.longtermplan.nhs.uk/ [accessed 2021-12-13]

XSL•FO
RenderX

2. Dobson J. Co-production helps ensure that new technology succeeds. BMJ 2019 Jul 25;41(6):l4833-l4850. [doi: 10.1136/bmj.l4833]
3. Palumbo R. Contextualizing co-production of health care: a systematic literature review. Int J Public Sector Manag 2016 Jan 11;29(1) [FREE Full text] [doi: 10.1108/ijpsm-07-2015-0125]
4. Topo P. Technology studies to meet the needs of people with dementia and their caregivers. J Appl Gerontol 2008 Oct 01;28(1):5-37 [FREE Full text] [doi: 10.1177/0733464808324019]
5. Martin A, Caon M, Adorni F, Andreoni G, Ascolese A, Atkinson S, et al. A mobile phone intervention to improve obesity-related health behaviors of adolescents across Europe: iterative co-design and feasibility study. JMIR Mhealth Uhealth 2020 Mar 02;8(3):e14118 [FREE Full text] [doi: 10.2196/14118] [Medline: 32130179]
6. Sanders E, Stappers P. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18 [FREE Full text] [doi: 10.1080/15710880701875068]
7. Steen M, Manschot M, Koning N. Benefits of co-design in service design projects. Int J Design 2011 Aug;5(2):53-60 [FREE Full text]
8. Cavalcanti L, Holden R, Karanam Y. Applying participatory design with dementia stakeholders: challenges and lessons learned from two projects. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare – Demos, Posters, Doctoral Colloquium. 2018 Presented at: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare – Demos, Posters, Doctoral Colloquium; May 21-24, 2018; New York, United States. [doi: 10.4108/eai.20-4-2018.2276362]
9. Kerkhof Y, Pelgrum-Keurhorst M, Mangiaracina F, Bergsma A, Vrauwdeunt G, Graff M, et al. User-participatory development of FindMyApps; a tool to help people with mild dementia find supportive apps for self-management and meaningful activities. Digit Health 2019;5:2055207618822942 [FREE Full text] [doi: 10.1177/2055207618822942] [Medline: 30944726]
10. Miah J, Dawes P, Edwards S, Leroi I, Starling B, Parsons S. Patient and public involvement in dementia research in the European Union: a scoping review. BMC Geriatr 2019 Aug 14;19(1):220 [FREE Full text] [doi: 10.1186/s12877-019-1217-9] [Medline: 31412788]
11. Pickett J, Murray M. Editorial: patient and public involvement in dementia research: setting new standards. Dementia (London) 2018 Nov;17(8):939-943 [FREE Full text] [doi: 10.1177/1471301218789290] [Medline: 30373464]
12. Morgan N, Grinbergs-Saull A, Murraty M. 'We can make our research meaningful': the impact of the Alzheimer's Society Research Network. Alzheimer's Society. URL: https://www.alzheimers.org.uk/sites/default/files/2018-04/Research%20Network%20Report%20low-res.pdf [accessed 2021-12-13]
13. Hudson JM, Pollux PM. The Cognitive Daisy - A novel method for recognising the cognitive status of older adults in residential care: Innovative Practice. Dementia (London) 2019 Jul;18(5):1948-1958 [FREE Full text] [doi: 10.1177/1471301216673918] [Medline: 27758959]
14. Van der Roest HG, Wenborn J, Pastink C, Dröes RM, Orrell M. Assistive technology for memory support in dementia. Cochrane Database Syst Rev 2017 Jun 11;6:CD009627 [FREE Full text] [doi: 10.1002/14651858.CD009627.pub2] [Medline: 28602027]
15. Bateman DR, Srinivas B, Emmett TW, Schleyer TK, Holden RJ, Hendrie HC, et al. Categorizing health outcomes and efficacy of mHealth apps for persons with cognitive impairment: a systematic review. J Med Internet Res 2017 Aug 30;19(8):e301 [FREE Full text] [doi: 10.2196/jmir.7814] [Medline: 28855146]
16. Andrews JA, Brown LJ, Hawley MS, Astell AJ. Older Adults' Perspectives on Using Digital Technology to Maintain Good Mental Health: Interactive Group Study. J Med Internet Res 2019 Feb 13;21(2):e11694 [FREE Full text] [doi: 10.2196/11694] [Medline: 30758292]
17. Hassan L, Swarbrick C, Sanders C, Parker A, Machin M, Tully MP, et al. Tea, talk and technology: patient and public involvement to improve connected health 'wearables' research in dementia. Res Involv Engagem 2017 Aug 01;3:12 [FREE Full text] [doi: 10.1186/s40900-017-0063-1] [Medline: 29062537]
18. May C, Hasher L, Stoltzfus E. Optimal time of day and the magnitude of age differences in memory. Psychol Sci 2017 Apr 25;4(5):326-330 [FREE Full text] [doi: 10.1111/j.1467-9280.1993.tb00573.x]
19. Mortamais M, Ash JA, Harrison J, Kaye J, Kramer J, Randolph C, et al. Detecting cognitive changes in preclinical Alzheimer's disease: a review of its feasibility. Alzheimers Dement 2017 Apr 01;13(4):468-492. [doi: 10.1016/j.jalz.2016.06.2365] [Medline: 27702618]
20. Menegardo CS, Friggi FA, Scardini JB, Rossi TS, Vieira TD, Tieppo A, et al. Sundown syndrome in patients with Alzheimer's disease dementia. Dement Neuropsychol 2019 Sep 01;13(4):469-474 [FREE Full text] [doi: 10.1590/1980-57642018dn13-040015] [Medline: 31844502]
21. Khachiyants N, Trinkle D, Son SJ, Kim KY. Sundown syndrome in persons with dementia: an update. Psychiatry Investig 2011 Dec 01;8(4):275-287 [FREE Full text] [doi: 10.4306/pi.2011.8.4.275] [Medline: 22216036]
22. Dementia assessment and diagnosis. National Institute for Health and Care Excellence. URL: https://pathways.nice.org.uk/pathways/dementia/dementia-assessment-and-diagnosis [accessed 2021-12-13]

XSL•FO
RenderX

23. Brown LJ, Adlam T, Hwang F, Khadra H, Maclean LM, Rudd B, et al. Computer-based tools for assessing micro-longitudinal patterns of cognitive function in older adults. Age (Dordr) 2016 Aug;38(4):335-350 [FREE Full text] [doi: 10.1007/s11357-016-9934-x] [Medline: 27473748]
24. Muniz-Terrera G, Watermeyer T, Danso S, Ritchie C. Mobile cognitive testing: opportunities for aging and neurodegeneration research in low- and middle-income countries. J Glob Health 2019 Dec;9(2):020313 [FREE Full text] [doi: 10.7189/jogh.09.020313] [Medline: 31893027]
25. O'Connor S. Co-designing technology with people with dementia and their carers: exploring user perspectives when co-creating a mobile health application. Int J Older People Nurs 2020 Sep;15(3):e12288. [doi: 10.1111/opn.12288] [Medline: 31837096]
26. Yousaf K, Mehmood Z, Awan IA, Saba T, Alharbey R, Qadah T, et al. A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD). Health Care Manag Sci 2020 Jun 01;23(2):287-309. [doi: 10.1007/s10729-019-09486-0] [Medline: 31218511]
27. Roberts M. Successful public health information system database integration projects: a qualitative study. Online J Public Health Inform 2018;10(2):e207 [FREE Full text] [doi: 10.5210/ojphi.v10i2.9221] [Medline: 30349625]
28. Raghu A, Praveen D, Peiris D, Tarassenko L, Clifford G. Engineering a mobile health tool for resource-poor settings to assess and manage cardiovascular disease risk: SMARThealth study. BMC Med Inform Decis Mak 2015 Apr 29;15:36 [FREE Full text] [doi: 10.1186/s12911-015-0148-4] [Medline: 25924825]
29. Tambuyzer E, Pieters G, Van Audenhove C. Patient involvement in mental health care: one size does not fit all. Health Expect 2014 Feb 01;17(1):138-150 [FREE Full text] [doi: 10.1111/j.1369-7625.2011.00743.x] [Medline: 22070468]
30. Dementia-friendly communities. Alzheimer's Society. URL: https://www.alzheimers.org.uk/get-involved/dementia-friendly-communities [accessed 2021-12-13]
31. Good practice for payment and recognition - things to consider. NIHR|INVOLVE. 2019. URL: https://www.invo.org.uk/good-practice-for-payment-and-recognition-things-to-consider/ [accessed 2021-08-01]
32. Turk A, Boylan A, Locock L. A researchers guide to patient public involvement. Healthtalk. URL: https://oxfordbrc.nihr.ac.uk/wp-content/uploads/2017/03/A-Researchers-Guide-to-PPI.pdf [accessed 2021-12-13]
33. Schwaber K. SCRUM development process. In: Business Object Design and Implementation. London: Springer; 1997.
34. Oomen S, Benny DW, Ademar A, Pascal R. How can scrum be successful? Competences of the scrum product owner. In: Proceedings of the 25th European Conference on Information Systems (ECIS). 2017 Jan 01 Presented at: Proceedings of the 25th European Conference on Information Systems (ECIS); Jun 5-10, 2017; Guimarães, Portugal.
35. Lazak M, Howieson D, Bigler E, Tranel D. Neuropsychological Assessment 5th edition. Oxford, UK: Oxford University Press; 2012.
36. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. Res Involv Engagem 2017 Aug 2;3:13 [FREE Full text] [doi: 10.1186/s40900-017-0062-2] [Medline: 29062538]
37. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
38. Guidance on co-producing a research project. NIHR. 2018 Apr 01. URL: https://www.invo.org.uk//wp-content//uploads//2019//04//Copro_Guidance_Feb19.pdf [accessed 2021-12-13]
39. Brown LJ, Dickinson T, Smith S, Brown Wilson C, Horne M, Torkington K, et al. Openness, inclusion and transparency in the practice of public involvement in research: a reflective exercise to develop best practice recommendations. Health Expect 2018 Apr;21(2):441-447 [FREE Full text] [doi: 10.1111/hex.12609] [Medline: 29105227]
40. Nan N, Harter D. Impact of budget and schedule pressure on software development cycle time and effort. IIEEE Trans Software Eng 2009 Sep 01;35(5):624-637. [doi: 10.1109/tse.2009.18]

Abbreviations

PPI: patient and public involvement

Edited by L Buis; submitted 30.10.20; peer-reviewed by HL Tam, M Verbiest, YL Leung; comments to author 01.12.20; revised version received 20.01.21; accepted 08.10.21; published 14.01.22.

Please cite as:
Fox S, Brown LJE, Antrobus S, Brough D, Drake RJ, Jury F, Leroi I, Parry-Jones AR, Machin M
Co-design of a Smartphone App for People Living With Dementia by Applying Agile, Iterative Co-design Principles: Development and Usability Study
JMIR Mhealth Uhealth 2022;10(1):e24483
URL: https://mhealth.jmir.org/2022/1/e24483
doi:10.2196/24483
PMID:35029539

XSL•FO
RenderX

©Sarah Fox, Laura J E Brown, Steven Antrobus, David Brough, Richard J Drake, Francine Jury, Iracema Leroi, Adrian R Parry-Jones, Matthew Machin. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 14.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study

Aditya Vaidyam[1], MS; John Halamka[2], MD; John Torous[1], MBI, MD

[1]Beth Israel Deaconess Medical Center, Boston, MA, United States

[2]Mayo Clinic, Rochester, MN, United States

Corresponding Author:
John Torous, MBI, MD
Beth Israel Deaconess Medical Center
330 Brrokline Avenue
Boston, MA, 02215
United States
Phone: 1 6176676700
Email: jtorous@bidmc.harvard.edu

Abstract

Background: There is a growing need for the integration of patient-generated health data (PGHD) into research and clinical care to enable personalized, preventive, and interactive care, but technical and organizational challenges, such as the lack of standards and easy-to-use tools, preclude the effective use of PGHD generated from consumer devices, such as smartphones and wearables.

Objective: This study outlines how we used mobile apps and semantic web standards such as HTTP 2.0, Representational State Transfer, JSON (JavaScript Object Notation), JSON Schema, Transport Layer Security (version 1.3), Advanced Encryption Standard-256, OpenAPI, HTML5, and Vega, in conjunction with patient and provider feedback to completely update a previous version of mindLAMP.

Methods: The Learn, Assess, Manage, and Prevent (LAMP) platform addresses the abovementioned challenges in enhancing clinical insight by supporting research, data analysis, and implementation efforts around PGHD as an open-source solution with freely accessible and shared code.

Results: With a simplified programming interface and novel data representation that captures additional metadata, the LAMP platform enables interoperability with existing Fast Healthcare Interoperability Resources–based health care systems as well as consumer wearables and services such as Apple HealthKit and Google Fit. The companion Cortex data analysis and machine learning toolkit offer robust support for artificial intelligence, behavioral feature extraction, interactive visualizations, and high-performance data processing through parallelization and vectorization techniques.

Conclusions: The LAMP platform incorporates feedback from patients and clinicians alongside a standards-based approach to address these needs and functions across a wide range of use cases through its customizable and flexible components. These range from simple survey-based research to international consortiums capturing multimodal data to simple delivery of mindfulness exercises through personalized, just-in-time adaptive interventions.

(JMIR Mhealth Uhealth 2022;10(1):e30557) doi:10.2196/30557

KEYWORDS

digital phenotyping; mHealth; apps; FHIR; digital health; health data; patient-generated health data; mobile health; smartphones; wearables; mobile apps; mental health, mobile phone

Introduction

Background

The medical field today is transitioning toward integrating patient-generated health data (PGHD) into clinical care to increase shared decision-making, coordination of care, patient safety, and clinical outcomes [1]. PGHD are central to this mission and are defined as data recorded or created by the patient or caregivers used to address health concerns. Examples include a longitudinal view of symptoms of patient status between clinic visits captured via an app or information related to daily adherence to treatment plans [1]. This may include a

XSL•FO
RenderX

daily step count captured from a smartphone, medication surveys administered on the smartphone, and sleep quality data measured via a wearable sensor. Given the ability of smartphones and wearables to collect a myriad of continuous multimodal data relevant to care, such as heart rate, sleep, steps, and more, tools and systems to harness and use this vast amount of automatically generated PGHD are a health care priority. One challenge that remains toward this goal is the lack of technical infrastructure and organizational capability to handle the intake of accurate and valid PGHD from patient-owned consumer devices. Such standards and tools supporting the effective and compatible integration of PGHD are needed to enhance clinical insight and support research and data analysis before PGHD can actually impact routine care.

Digital Phenotyping+

The need for apps that can not only capture but also use and integrate PGHD is clear. The use of commercially available wearable technology for the acquisition of PGHD has seen a recent uptick owing to the effects of the COVID-19 pandemic and growing demand for telehealth services [2,3]. Today, over 80% of Americans own a smartphone device [4] and over 20% own a wearable device [5], reinforcing the potential of PGHD to improve clinical outcomes. For example, the Apple Watch today retails at US $350 with onboard nonmedical grade electrocardiography and oxygen saturation sensors that continuously measure and record data. Modern smartphones are also equipped with numerous sensors that generate a high volume of potentially clinically significant PGHD that could enable a better real-time understanding of cognition, mobility, sociability, and more through techniques, such as digital phenotyping and ecologic momentary assessment.

Digital phenotyping [6] is the construction of an individual-level phenotype using data collected from smartphones or wearable devices actively via user interaction (eg, surveys), or passively without user input (eg, sensors, such as an accelerometer). Although there are many digital phenotyping tools and systems used in health care and research contexts, a recent review identified nearly 50 of them [7], and few offer an integrated and standardized approach to analyze and respond to clinically actionable patient-generated data. Existing tools are primarily closed systems or consist of only a single app with little flexibility or customizability [7]. In a recent review, 85% of existing solutions supported active and passive sensing but only 33% supported clinical assessment, 30% supported predictive modeling of patient data, and 24% supported app-delivered interventions [7]. Furthermore, only 35% of the existing solutions showed a patient- or clinician-facing user interface [7]. However, the combination of all these features is necessary to meet the diverse needs of research and care delivery. A search outside the research literature and instead, across mental health smartphone apps in commercial marketplaces reported that only 1.1% supported sensors [8], suggesting that many research tools do not translate into accessible tools for patient or clinician use. Although diverse functionality and innovation continue to exist across the entire app space, we have argued that there is a need for multiple uses of the same app, instead of using multiple apps in a fragmented manner, toward better supporting clinical research, integration, and implementation [9].

Challenges in Integration

Creating PGHD tools that use sensor and digital phenotyping tools in a more patient- and health system–centric manner is a common goal, but it remains challenging to achieve. Despite the prevalence of existing electronic medical record standards and tools, a 2019 review on the integration of PGHD into clinical practice, “integration [...] was extremely limited, and decision support capabilities were for the most part basic” [10]. The most widely adopted medical record standards initiative that can be used to link PGHD to medical records is Fast Healthcare Interoperability Resources (FHIR), led by the Health Level 7 organization [11], which is now adopted by many major health care systems and industry partners, including Apple, Google, Amazon, Microsoft, and others [12]. Its companion projects SMART (Specific, Measurable, Achievable, Realistic, and Timely) [13] and SMART Markers [14] build on FHIR and enable integration of third-party modules into medical record systems, including patient mobile devices and sensors.

However, the FHIR ecosystem alone does not address a number of concerns specific to the integration of PGHD into clinical systems. FHIR and the current data interoperability standard (United States Core Data for Interoperability) [15] were not developed for continuous high-velocity data, and its implementation in the health care ecosystem today is primarily read-only, although its data gathering and write-back ability continues to evolve. For example, the FHIR core does not allow for semantic equivalence of data that can be used to automate data matching or harmonization. As a result, it is not possible to work with both cognitive assessment scores and mobility or sociability metrics using the same analysis pipeline. This increases the effort required and time taken to work with PGHD, as clinicians or researchers must first preprocess data of different semantic types individually before being able to work with a data set as a whole. Although R4 extensions, such as the mCODE core cancer model [16] are becoming a new way to expand FHIR’s supported vocabulary, they are still early in evolution and adoption. Today, the inability to standardize terminologies across interconnected systems, such as through a *data dictionary*, impedes effective export and analysis of different types of data from different data sources using the FHIR ecosystem [17,18]. These challenges preclude the adoption of FHIR as a PGHD-first standard for clinical and research use cases.

Thus, there remains a need for a flexible, interoperable, and extensible platform that enables the effective use of PGHD through widely accepted standards for both clinical and research needs. In this paper, we present a potential solution for the robust and effective acquisition and integration of PGHD into research and clinical care with tangible examples and open-source code.

Methods

Overview

To integrate PGHD into research and clinical care needs, our team has designed and developed the Learn, Assess, Manage, and Prevent (LAMP) platform that encompasses a robust set of protocols, standards, tools, and apps. Our team initially developed the mindLAMP smartphone app [19] as part of the

XSL•FO
RenderX

initial version of the LAMP platform. In this paper, we review a rearchitected and redeveloped platform comprising new frontend, backend, and analysis components to support PGHD, patient-centric care, and actionable digital phenotyping. This new platform, distinct from its predecessor, includes features such as customizable and schedulable activities, sensor data collection and analysis, messaging support with the care team, and more, available across modern web browsers and smartphone operating systems. The design and development of the platform was approached from both a patient- and clinician-focused approach as well as a semantic standards–based approach.

Patient- and Clinician-Led Design

The LAMP platform was designed and developed with continuous feedback from patients with serious mental illnesses and clinicians. Through a patient advisory panel, focus groups [20-22], clinical use, and feedback from a global consortium of users, mindLAMP has been co-designed iteratively with updates reflecting expanding ideas for its role. User input informed the adaptability, flexibility, and customizability of the LAMP platform, which resulted in a new user interface compared with the previous version, as shown in Figure 1. We established a formal system to enable anyone to suggest improvements, report bugs, and assess new features to ensure that all could partake in the iterative design rounds. This process also influenced aspects of the user experience, such as making mindLAMP available in multiple languages (English, Spanish, and Hindi) and designing to ensure easy addition of more. Key examples of patient feedback and the design outcomes they influenced are provided in Table 1.

Figure 1. (A) The home screen interface of the original (version 1 with red border) mindLAMP app; (B) the new and improved home screen interface of the (version 2 with green border) mindLAMP app that incorporates multiple activities and schedules into a heads-up tab called the Feed; (C) the Learn, Assess, Manage, and Prevent tabs group embedded activities together with helpful tooltips and icons, with new activity types (eg, tips, meditation); (D) returning data and insight to users was considered a priority in the redesigned user interface and thus more advanced charting tools were integrated into the Prevent tab, accessible to both patients and clinicians.

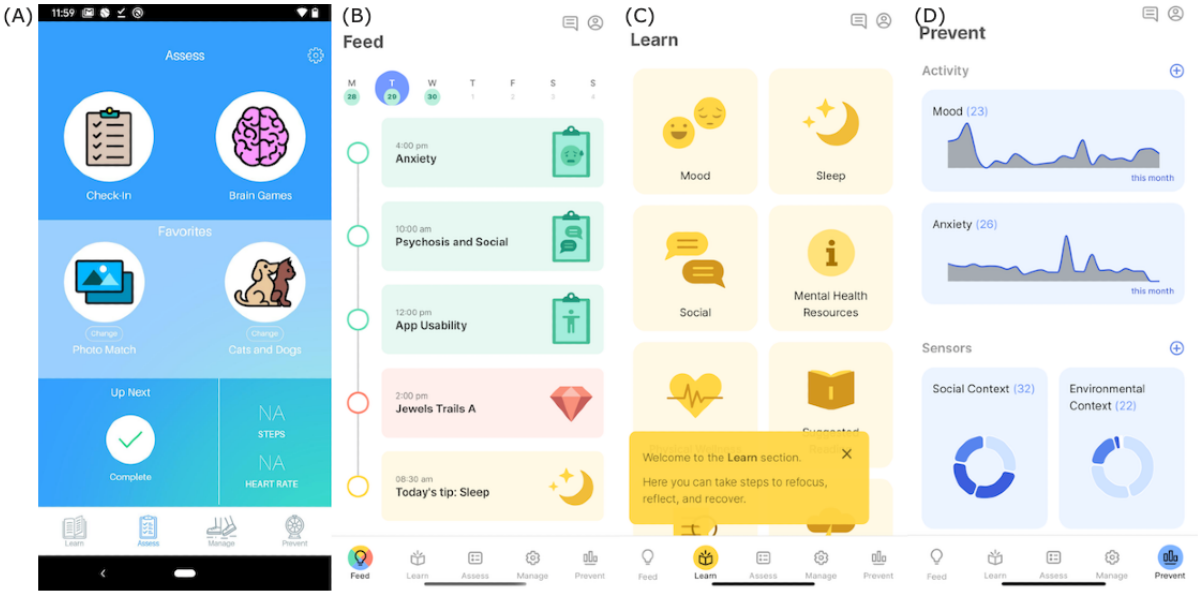
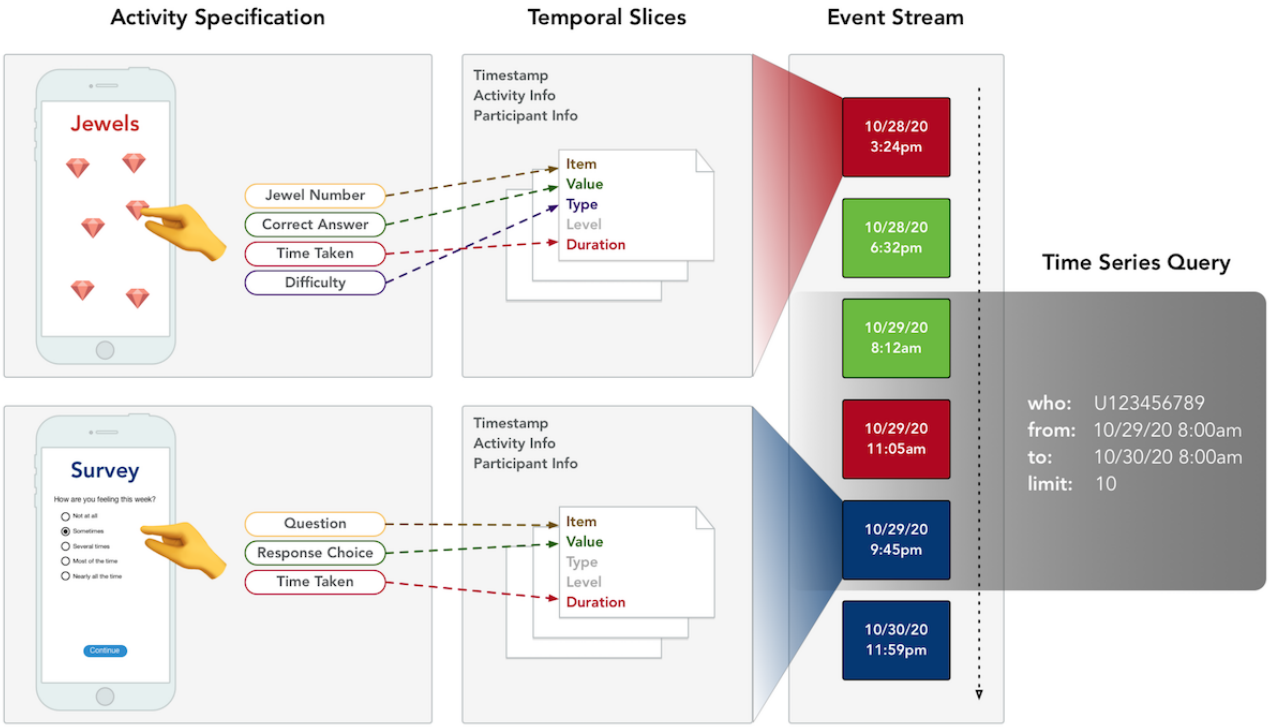


Table 1. Selected examples of patient feedback driving significant changes in the user experience and overall architecture of the Learn, Assess, Manage, and Prevent platform. Semantic (technical) standards–based approach.

Sample patient feedback	Outcome
"Apps like Facebook or Amazon are clear where I am lost in a sea of people or items and that is generally accepted, but with apps in the health space, who is involved, which institution is involved, level of comfort with the individuals and what data is collected, all of these factors are carefully calculated when I make a decision to join a study like this—an establishment of trust is crucial. Apps that track people incur a level of suspicion that changes between people, from none at all to a lot, perhaps depending on level of illness."	The Learn, Assess, Manage, and Prevent platform was rebuilt around an open-source collaborative environment supported by the consortium; all development and data handling processes are disclosed in the privacy policy, and significant backend changes were made to patient data equity and ownership.
"Yeah, because I don't see any apps out there these days that help people with psychosis and when they're getting sicker. It just seems...they just don't help with certain things. This gives you control to go get help if somebody needs it. It's like, the good thing about this app is, it's getting the right information and it's sending you somewhere, it's almost as if you could go to the therapist with this information! You don't want an app that's just one sided [and siloed off from the therapist or delivery of care]."	Additional types of activities were added to the mindLAMP app, including tips (Figure 2), meditation, and other informational and management tools; each of these activities captures metadata during patient use that can be interpreted and incorporated into a clinical encounter.
"mindLAMP is a tool for me to get better: I want to know if I'm making progress and when, what am I deficient in, how am I deficient, and how to improve on it; that is, as a metrics-driven person."	The smaller heads-up summaries originally found in the first version of the mindLAMP app (Figure 2) were updated and expanded into an entire tab (Figure 2), providing more insight and customization into patient data.

XSL•FO
RenderX

Figure 2. Flow of the data collection process from native app to backend: (1) an activity specification describes the types of interactive elements available in the mindLAMP app, along with their possible configuration parameters; (2) when participants interact with a configured and scheduled activity (such as a mood survey based on the survey specification), all metadata and data from the interaction session is integrated into a single unit of patient-generated health data called events; (3) events are then submitted to the backend in real time as part of a continuously generated stream of patient-generated health data; and (4) clinicians and researchers are able to perform continuously updating queries on the data with their desired parameters.



In addition to a patient- and clinician-centric design approach, the LAMP platform was also architected with a semantic standards–based approach, considering technical best practices for future proofing and security or compliance across health care systems. The open standards listed in Table 2 were chosen specifically to foster an open ecosystem around the platform. For example, the platform's programming interface adopts a *repository* model to store and configure patient-facing instruments, each with its own embedded user interface. These embedded user interfaces were developed using common and widely adopted web standards indicated in Table 2 (HTML5.0, Cascading Style Sheets 3.0, and ECMAScript 6). As a patient begins an interaction session, this embedded code is securely sandboxed by the mindLAMP user interface both within the smartphone app and the patient-facing web dashboard. In addition to providing a standard schematic of all structured documents encountered and processed in the LAMP platform using OpenAPI [23], the JSON Schema [24] data markup standard is used to provide developers of these interactive patient-facing instrument configurability and extensibility. With little required skill or upfront effort, developers can use the platform's software development kit to create instruments with completely customizable user experiences that are then tuned and customized by clinicians for individual patients or by research coordinators for studies spanning many patients.

XSL•FO
RenderX

Table 2. Adopted semantic web standards, their use rationale, and implementation details.

Standard	Description	Use	Reason chosen
HTTP 2 [22]	Ubiquitous web standard that declares and defines the semantics of client-server communication with a rich and readily available debugging and implementation toolset and ecosystem not available for custom binary protocols	Implemented by core programming libraries and the backend	In contrast to TCP[a]-based binary data protocols requiring specialized tooling to access and work with data, almost all systems and tools are able to interact with web standards through the HTTP protocol.
REST[b] [25]	Ubiquitous lightweight HTTP-based web standard that defines systems logically through accessibility and manipulation of remote resources instead of invocation of remote functions	Implemented by core programming libraries and the backend	In contrast to a custom implementation of remote function invocation that would require custom programming libraries to interface with, most web systems are able to interact with REST-based resources in a logical manner. In the absence of developer knowledge or pre-existing tools, it remains possible to communicate with RESTful systems.
JSON[c] [26]	Ubiquitous web standard that supports structured (as opposed to tabular, ie, CSV files) formatting and markup of data using strict data types	Implemented across all components in the platform	In contrast to encoded binary data formats requiring specialized tooling to interpret and work with data, most programming environments support the JSON standard.
TLS[d] version 1.3 [27]	Ubiquitous web standard that enables encryption of data in transit between client and server	Implemented by core modules and programming libraries, used by all components in the platform	No alternative
AES-256[e] [28]	Ubiquitous cryptographic standard that enables encryption of data at REST (on disk) by a database	Implemented by the backend and mandated by the backend and deployment configuration for the database within which data shall be stored	No alternative
JSON Schema [24]	Web standard that describes JSON-encoded data and metadata through ahead-of-time specification of a universally agreed upon schematic, as opposed to inline schema provided only at runtime	Implemented by the backend and used by the frontend	Although binary protocols require a predetermined strict schema to format the data, JSON does not. JSON Schema provides ahead-of-time resolution of the contents of a data payload and can be used to validate and harmonize data as well.
OpenAPI [23]	Web standard that describes REST-based web services and metadata through ahead-of-time specification of a universally agreed upon schematic	Implemented by the backend and core programming libraries and used by the frontend	In contrast to writing programming libraries and testing or validation tools, the generation of these tools and packages by the OpenAPI ecosystem increases productivity.
HTML5 [29]	Ubiquitous web standard that makes it possible to securely embed custom user interfaces backing patient-facing activities	Implemented by the frontend and all patient-facing activities, with wide support for CSS3[f] and JavaScript 2016 (ES6[g])	No alternative
Vega [30]	Visualization grammar standard that encodes charts and graphs as JSON documents that are then rendered and viewed interactively by apps	Implemented using HTML5. Implemented by the frontend and Cortex analysis code	In contrast to static images and handwritten analysis code, the ability to declaratively generate interactive real-time charts through an embedded query reduces data science and clinician effort and fatigue.

[a]TCP: transmission control protocol.

[b]REST: Representational State Transfer.

[c]JSON: JavaScript Object Notation.

[d]TLS: Transport Layer Security.

[e]AES-256: Advanced Encryption Standard.

[f]CSS: Cascading Style Sheets.

[g]ES: ECMAScript.

LAMP Platform

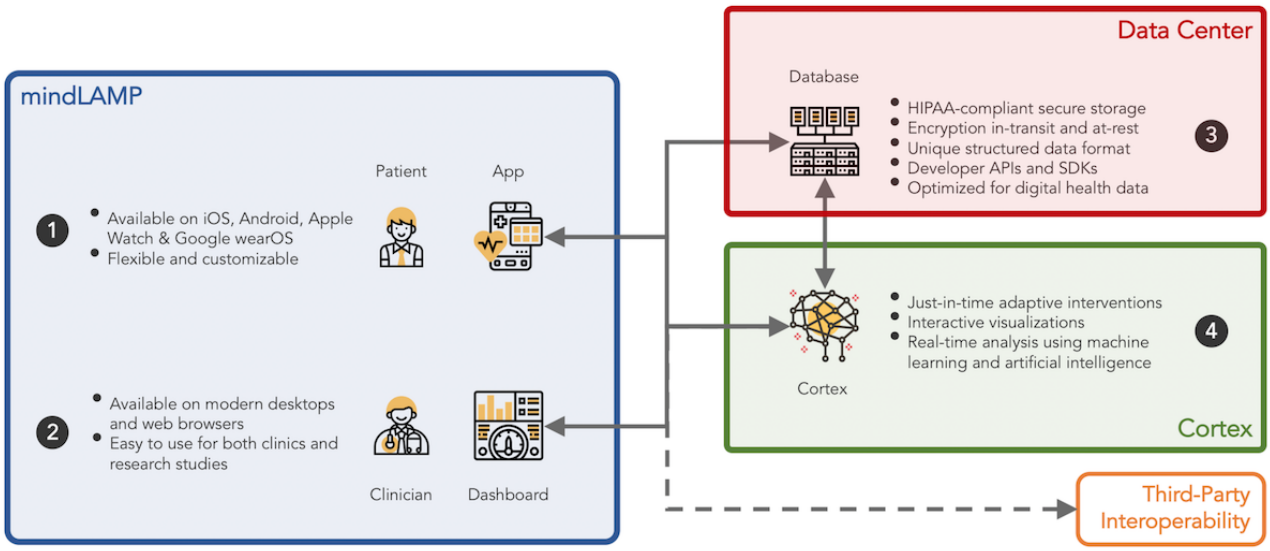
The LAMP platform is a customizable clinical care management and neuropsychiatric research platform designed around PGHD, as detailed in Figure 3. It comprises numerous essential features, such as customizable clinician-defined activities (eg, surveys, breathing exercises, journaling, and cognitive tests), collection and analysis of mobile and wearable sensor data, push notification scheduling, care team–centric conversations, just-in-time adaptive interventions, prebuilt featurization, visualization, or analysis pipelines, and a companion integrated development environment (IDE). The LAMP platform is available for use across any modern desktop web browser as well as recent versions of iPhone operating system and Android

XSL•FO
RenderX

through the mindLAMP app available on the Apple and Google app stores, as shown in Figure 4. The backend is deployable using enterprise-standard orchestration tools (Docker and Kubernetes [31]) and has already been deployed across several health care systems and is used today by patients, clinicians, and researchers. The companion Cortex data analysis toolkit integrates tightly across the platform to provide a unified processing pipeline for secondary active and passive data features (measurable behavioral characteristics extracted from raw data), interactive visualizations, and the generation of targeted and automated adaptive interventions. The IDE is bundled with support for the widely adopted Python, R, and JavaScript programming languages and built atop Jupyter Notebooks and Visual Studio Code for collaborative data analysis.

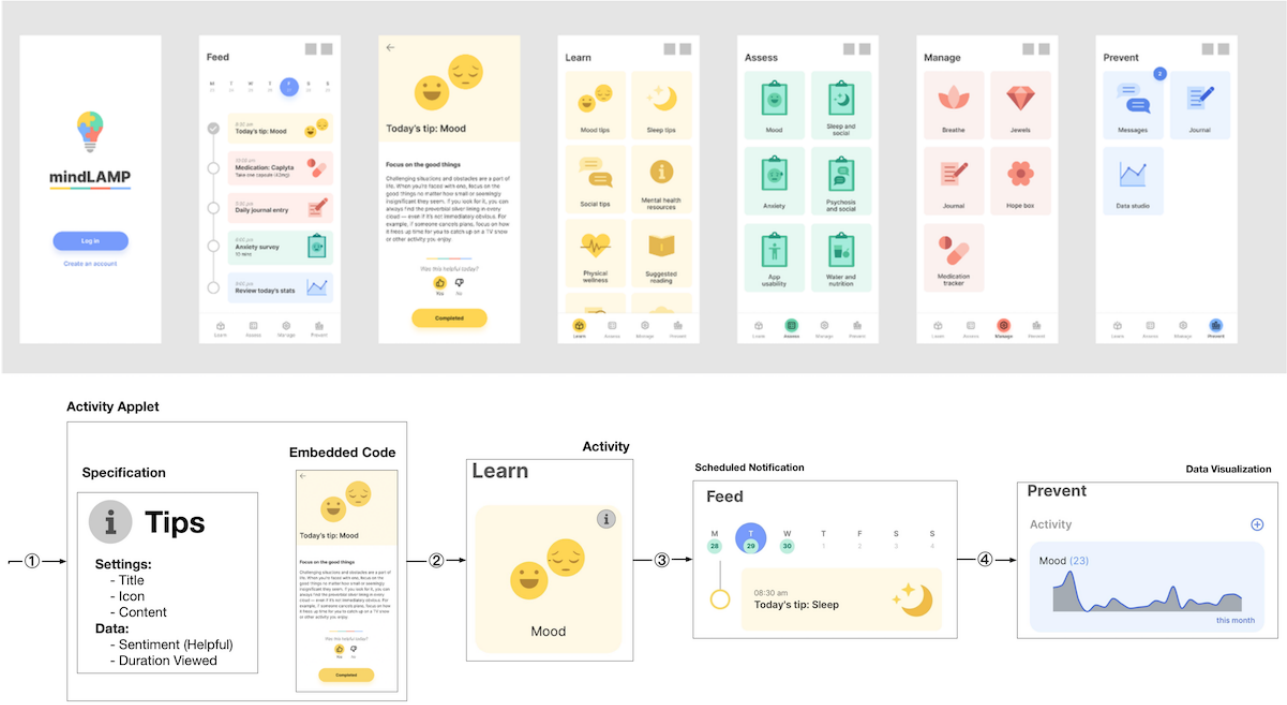
The LAMP platform is designed to be customizable to fit a wide range of use cases and requirements, eliminating the need for multiple apps hosting only a set of fixed, immutable content as well as the research concern of proprietary data formats and closed-source commercial analysis software. It also securely enables data interoperability and extensibility, avoiding the issue of *data silos* without external access of collected data for clinicians or patients. It integrates into existing hospital organization structure and is not limited to the *sandbox* on an individual's smartphone, allowing the caregivers and patients to coexist on the same platform. These features combined allow the LAMP platform to engage the care team through interactive clinical decision support with adaptive responses to incoming PGHD. Where a self-contained app must focus on solving individual problems for specific stakeholders, the LAMP platform focuses on broader challenges around linking people with data and data to teams of interconnected stakeholders, from patients and clinicians and family members and the care team to administrators and research coordinators. For these reasons, the LAMP platform supports *digital phenotyping+*, the plus symbol indicates the ability to return and share PGHD with the individuals from which it is collected in a secure and ethical manner and the ability to integrate that data into a machine learning pipeline or other clinical decision support algorithms.

Figure 3. The Learn, Assess, Manage, and Prevent platform consists of three major components: (1) mindLAMP, the patient- and clinician-facing app and web dashboard; (2) Data center, providing secure storage and access to data; and (3) Cortex, the data analysis toolkit that enables adaptive interventions and interactive visualizations. API: application programming interface; HIPAA: Health Insurance Portability and Accountability Act; iOS: iPhone operating system; SDK: software development kit.



XSL•FO
RenderX

Figure 4. Screenshots of the home screen as viewed by a patient in the mindLAMP app on a smartphone device; system administrators install applets such as Survey, Breathe, or Tips that clinicians and researchers are able to configure and schedule so that the participants interact with the app and produce data and metadata.



Results

Overview

The rearchitected LAMP platform addresses the integration of PGHD into existing systems through a simplified extensible programming interface (application programming interface [API]) and an internal data representation that interlinks raw data and metadata with descriptive schema. The Cortex data analysis toolkit obviates the need for custom preprocessing or harmonization of disparate sources of data and removes barriers between the real-time collection of PGHD and subsequent featurization, analysis, or visualization. We present the process results and examples below but do not offer a hypothesis in line with other papers exploring informatics systems created for use in clinical care and research.

Integration to Existing Systems

To enable robust data analysis, adaptive interventions, and interoperability with a broad range of health care systems and services, the platform's data repository and programming interface are based upon a concise semantic FHIR-compatible API. The platform's API provides both predefined and pluggable schematics for patient-facing instruments, such as surveys and cognitive tests as well as for mobile and wearable sensors. The platform's backend validates and harmonizes patient data upon receipt, retaining lossless FHIR compatibility in the process. The platform provides a facility to query and transform data into FHIR-compatible *bundle* and *resource* types, in addition to other domain-specific tabular or structured data formats. Table 2 lists the clinical, regulatory, and software standards implemented and supported by the LAMP platform.

The LAMP platform's internal data representation provides a simplified abstraction around PGHD in comparison with FHIR. Fundamentally, FHIR adapts a message and document-based exchange programming interface atop the representational exchange state transfer web standard protocol. The FHIR data structures (Figure 5) consist of over 90 *modules* for clinical use, insurance, billing, and other use based on the concept of *resources*, with each resource containing some raw data, metadata, a schema identifier, and a human-readable representation of the raw data. The schema identifier is used to reference how the data should be interpreted by a compatible system or machine. As the raw data contained within 2 resources of the same schema type may differ (eg, the use of the *observation* data type to represent both blood pressure and depression assessments results), data processing cannot be standardized across different data types. By organizing and accessing PGHD separately from generalized repositories of data, such as electronic health record systems using the FHIR API, common and shared analysis methods and processing tools that are standardized across such various data types can be used by clinicians and researchers.

XSL•FO
RenderX

Figure 5. The structured list of all supported Fast Healthcare Interoperability Resources core resources; the structured list of all Learn, Assess, Manage, and Prevent platform resources.

The LAMP platform declares only 12 core PGHD-centric resources (Figure 5) that remain focused on clinical and research use. The *Researcher* and *Study* resource types group together sets of *participants* as well as the *activities* and *sensors* they are able to interact with or collect data from. Upon data collection, *ActivityEvent* and *SensorEvent* describe and link the recorded data to its metadata and any specific customized parameters. The *Credential* resource provides security access controls to any of the aforementioned resources, and the *Tag* resource provides support for integration, extensibility, and backward compatibility. The semantic context of any recorded data and metadata is described by *ActivitySpec*, *SensorSpec*, and *TagSpec*.

XSL•FO
RenderX

Internal Data Representation

Understanding the need for seamless integration into existing health care systems, software, and services, the platform exposes its internal data representation through the LAMP protocol, a programming interface that enables integration with third-party services. For example, integration with Google Fit wearable devices that also implement this same *push-based model*, is possible by signing up on the Fitbit developer portal and connecting the data output of the Fitbit programming interface to the data input in the programming interface provided by the LAMP platform. In another example of seamless integration, clinicians and researchers can use automated scripts to synchronize data between the mindLAMP app and their existing record-keeping systems. Data can be proactively fetched and stored securely, and users of the platform are notified of any or all data from a particular patient using *subscriptions*, regardless of whether the data were generated by the mindLAMP app or a third-party data source.

The extensibility and flexibility of patient-facing instruments in the LAMP platform rely on the unique data structure and functionality provided by the LAMP protocol as shown in Figure 2. Each activity with which patients are able to interact is defined and encapsulated in an *activity specification* that contains the program code written using web-compatible standards, along with descriptors of the required input configuration and output data. When a patient begins an interactive session with any activity, session-wide metadata regarding *who*, *what*, and *when* are recorded. Each tap of the screen within the activity is then automatically validated and converted into a standardized data format called a *temporal slice*. When the user completes the interactive session, all the *temporal slices* are packaged into chronologically ordered events indexed under the patient's identifier as a stream of continuously generated data. The data analyst is then able to query these data at any desired temporal resolution (eg, 1 millisecond, 1 day, and 1 year) and filter by the type of activity (eg, mood survey, anxiety survey, trails-making test, and meditation). The query can be mutated using transformation logic executed by the backend and subscribed such that newly uploaded data matching the query is reported in real time to the data analyst. This query framework can be used to better understand how participants use and engage with the activities available to them as part of the study, for example, by extracting a real-time metric of duration spent meditating in the app per participant.

As depicted in Figure 4, the flow of activity specifications to configured activities to their generated PGHD facilitates patient interaction with any kind of interactive web media, from static text for tips, to video content for learning modules, or audio content for breathing exercises. Instruments and their data can be monitored and maintained organization-wide for compliance and conformance. Multimedia Appendix 1 lists the sources of active and passive data currently available within the mindLAMP app and their data sources and types. This novel data organization and structure supported by the platform enables unification and harmonization of these different data types, with both backward compatibility to data types from legacy systems and future compatibility for data types for systems that are not yet available.

Data Analysis With Cortex

The same pipeline operates on both active and passive data, unifying the conceptual model for PGHD processing and obviating the need for individual analyses tied to custom code for specific sensor types across various devices. Sensor data are therefore subject to additional harmonization to account for the various differences in functionality and recording between Apple and Android devices. For example, accelerometer measurements taken on Apple devices are measured in units of gravity (G) with a frame of reference experiencing −1 G in the downward-facing axis, whereas measurements on Android are measured in meters per second square (m/s^2) without a frame of reference provided. As the platform automatically applies this harmonization step, the data analysis code does not require an intrinsic understanding of the source of the data. Samples of sensor data after harmonization are shown in Figure 6. Furthermore, in addition to raw sensors on smartphones or wearable devices, processed Apple HealthKit and Google Fit sensor data, such as activity recognition or heart rate variability, are available to the LAMP platform.

The Cortex data analysis toolkit further simplifies the extraction of passive data features as listed in Multimedia Appendix 1, with an example shown in Figure 7. Cortex provides prebuilt, parallelized and vectorized workflows in Python for PGHD extraction and featurization that operate across large data sets to generate interactive visualizations for the mindLAMP dashboard using the Vega visualization grammar (as listed in Table 2). It obviates the need to work directly with the LAMP protocol, allowing data scientists to reason about live actionable structured data entirely as data frames within their programming environment of choice. Through the vectorization of array operations and parallelization of function calls, Cortex is able to target high performance and cost-effectiveness, while maintaining data security and policy compliance. A sample execution plan for a particular analysis involving the GPS data is shown in Figure 8. The modular nature of PGHD captured by mindLAMP allows for personalization and creation of new digital biomarkers and analysis without the need for additional coding. Furthermore, the companion IDE manager abstracts away log-in and security issues by securely injecting an authenticated connection to the server into Cortex and the resulting analysis notebooks.

XSL•FO
RenderX

Figure 6. Samples of data from selected sensors in the mindLAMP app for a sample patient. Total duration (in seconds) spent in calls per day; cumulative number of steps taken per hour during a 24-hour rolling window; number of times the device's screen was turned on per day; number of unique nearby devices (Wi-Fi or Bluetooth) encountered per day.

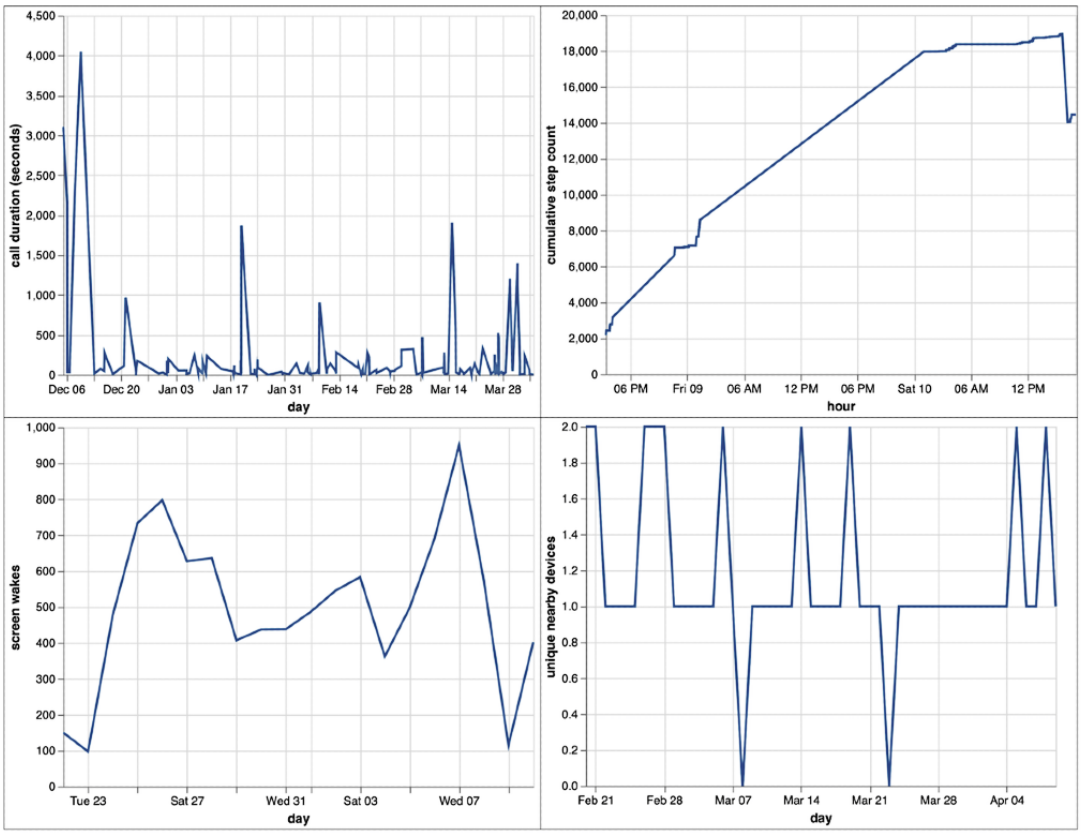
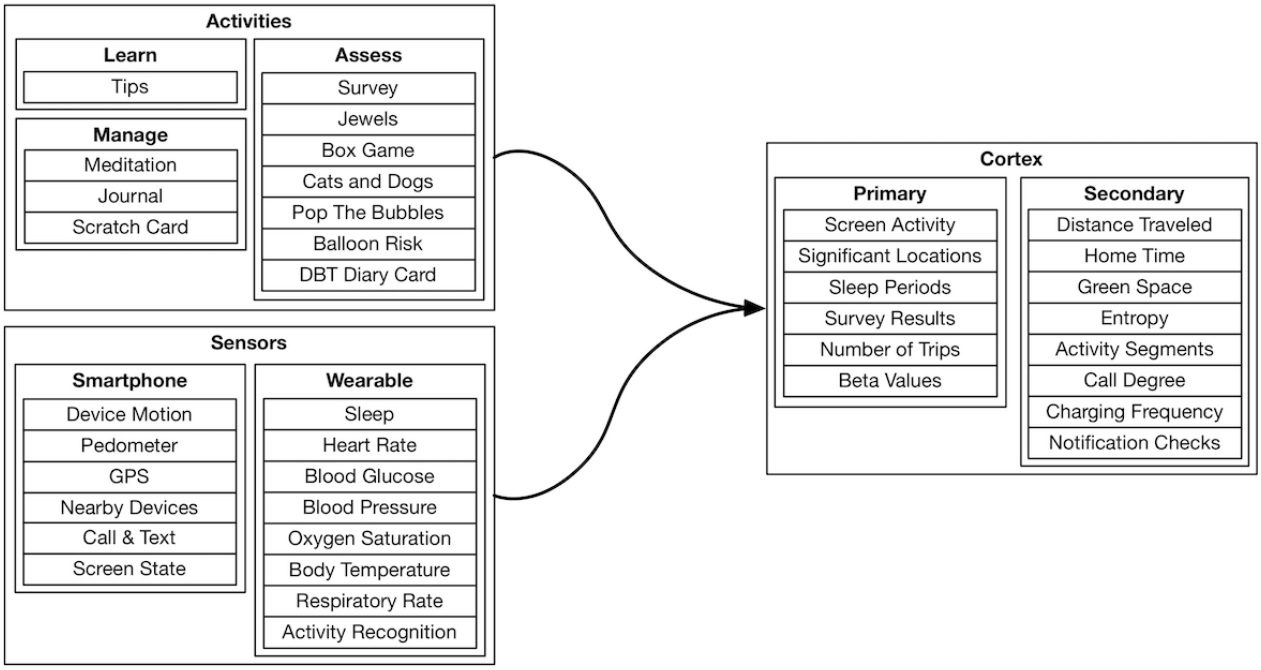
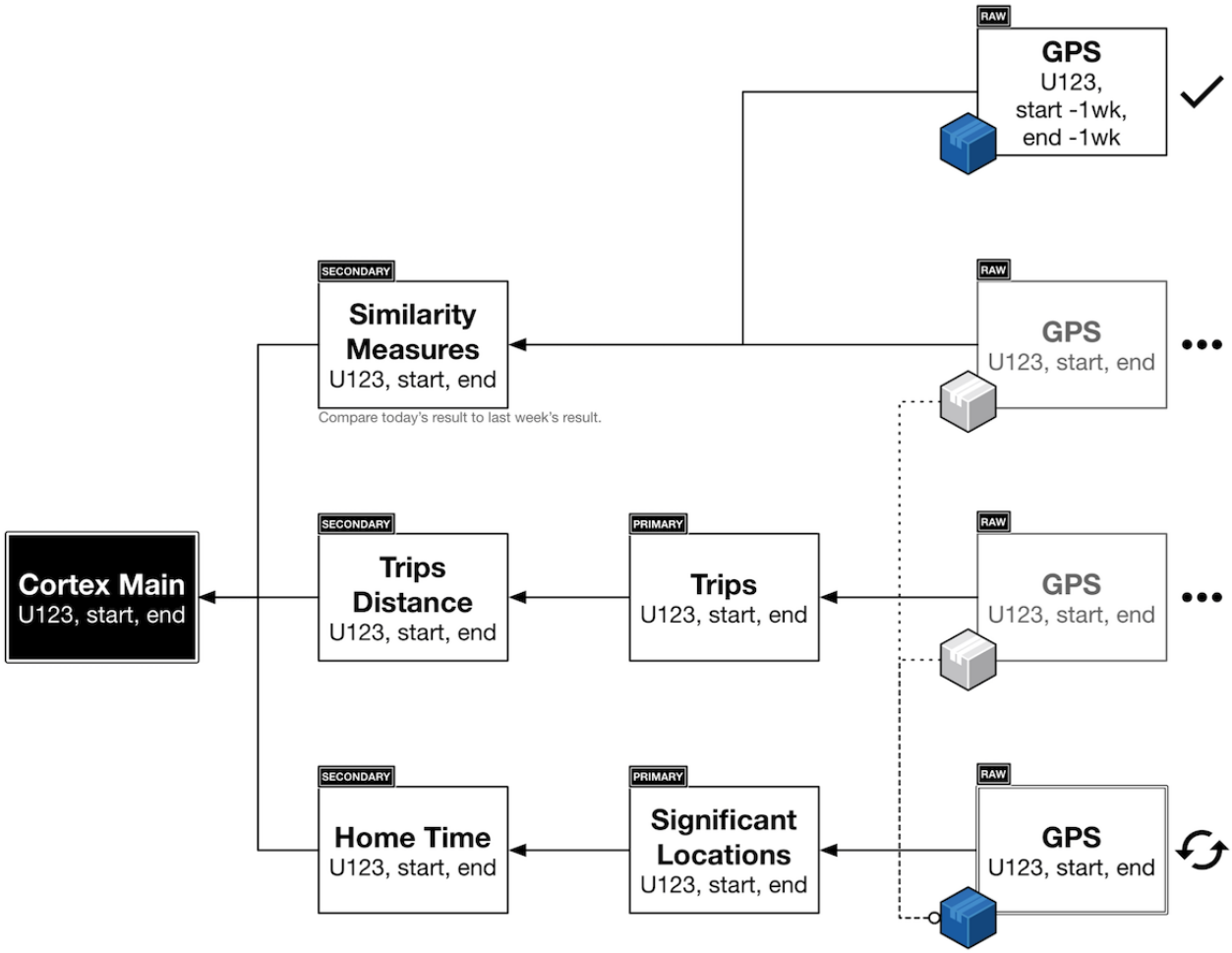


Figure 7. A visual representation of the various categories of activity and sensor data type features using standardized functions as part of the Cortex data analysis toolkit; shown as part of Cortex is the distinction between the primary and secondary feature types, where secondary features are composed of primary features as opposed to raw patient-generated health data. Availability of wearable sensors depends on the device type used and supported application programming interface; Apple Watch (HealthKit) sensors are shown here. DBT: dialectical behavioral therapy.



XSL•FO
RenderX

Figure 8. A sample execution plan for Cortex around an example using only geolocation data (1) the clinician or researcher creates an aggregate operation; (2) Cortex transparently interposes the correct feature layers by creating a dependency graph of data and executes each atomic operation (ie, independent of external variables) in the order it computes to be most efficient; (3) any raw sensor data are transparently cached during execution; and (4) as multiple operations require the same raw sensor data, Cortex blocks their execution until the cached data becomes available, to avoid duplicate downloads, wasted computation, and oversaturation of network bandwidth.



Discussion

Principal Findings

Research and clinical needs in digital medicine are evolving to use PGHD approaches to understand patient behavior and symptomatology [3]. To this end, by optimizing the system architecture for data throughput and substantial database write-loads, the LAMP platform supports high-performance data collection and real-time data analysis to enable, for example, larger machine learning models or just-in-time adaptive interventions that can leverage PGHD into actionable insights for patients and clinicians alike.

Efficient Collection and Configuration

Among the various approaches to data collection adopted in digital medicine, the *pull-based model* [13,14] shown in Figure 9, requires patients to activate a request and upload data from their mobile devices. This request can be scheduled and authorized. An example of the pull-based model is that during a clinical encounter, the clinician would use a portal to request data collection from the patient's device for a period of 1 day; the patient would then have to approve this request in their smartphone app before data collection can begin. During the next clinical encounter, the clinician would be able to interpret the data in potentially meaningful ways.

The LAMP platform, however, adopts a *push-based model* shown in Figure 9, where, in contrast, clinicians or research coordinators configure and schedule activities for patients to use and sensors from which measurements should be passively recorded ahead of time. The patients' devices receive a configuration request that activates data collection in the background. As it is collected, the data are uploaded (*pushed*) to the back end periodically, available for processing and clinical insight ahead of time. This push-based approach reduces latency from collection of PGHD to the usability of that PGHD, for example, as real-time alerts in the context of research studies (Figures 10 and 11), or toward clinical decision-making with custom rules and alerts. It is important that clinicians or research coordinators communicate clearly and establish consent with the patient or subject about the various types of data being collected and the frequency of the data push.

XSL•FO
RenderX

Figure 9. A pull-based model, in which clinicians must schedule data to be pulled from the device periodically versus a push-based model, in which clinicians preconfigure various sensors on the device to collect and push data to the server in real time.

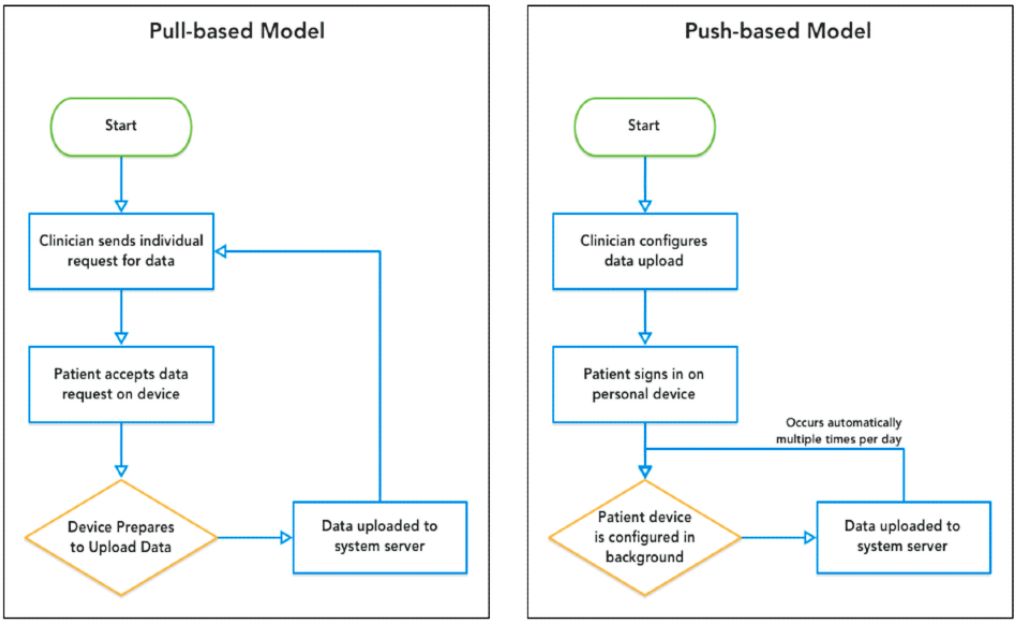
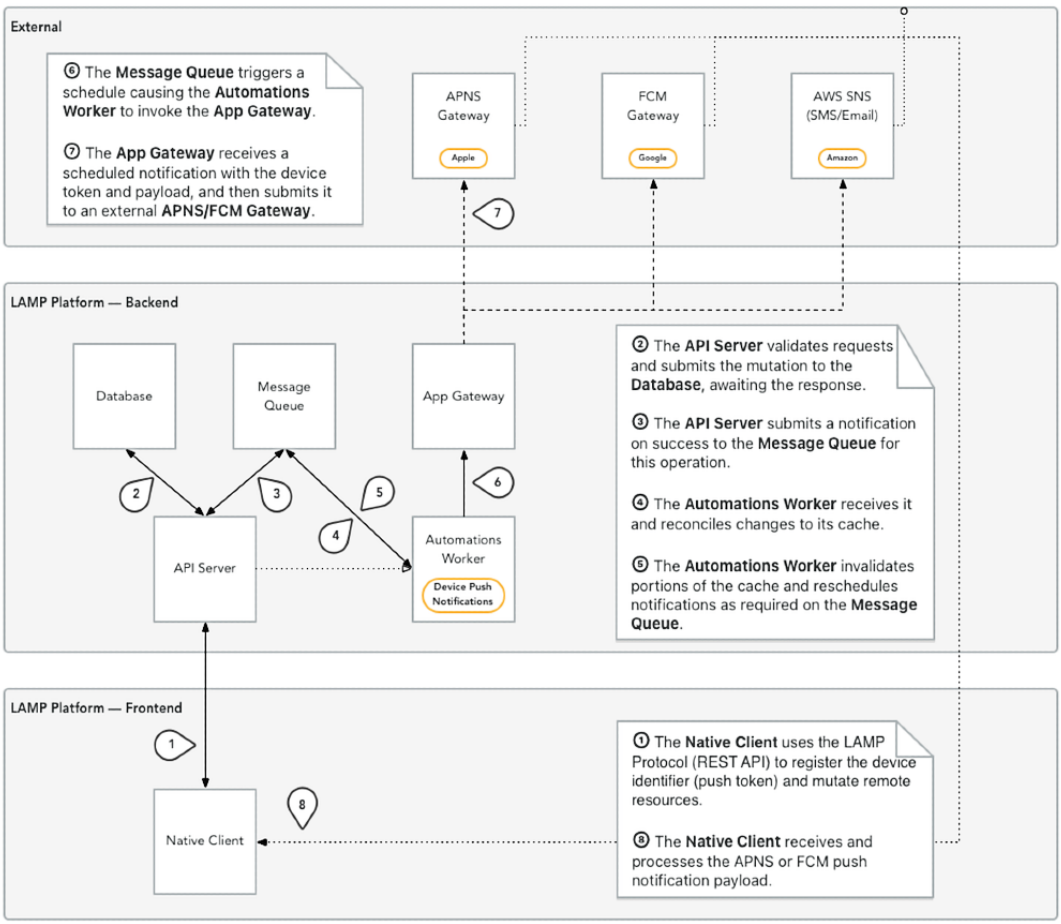
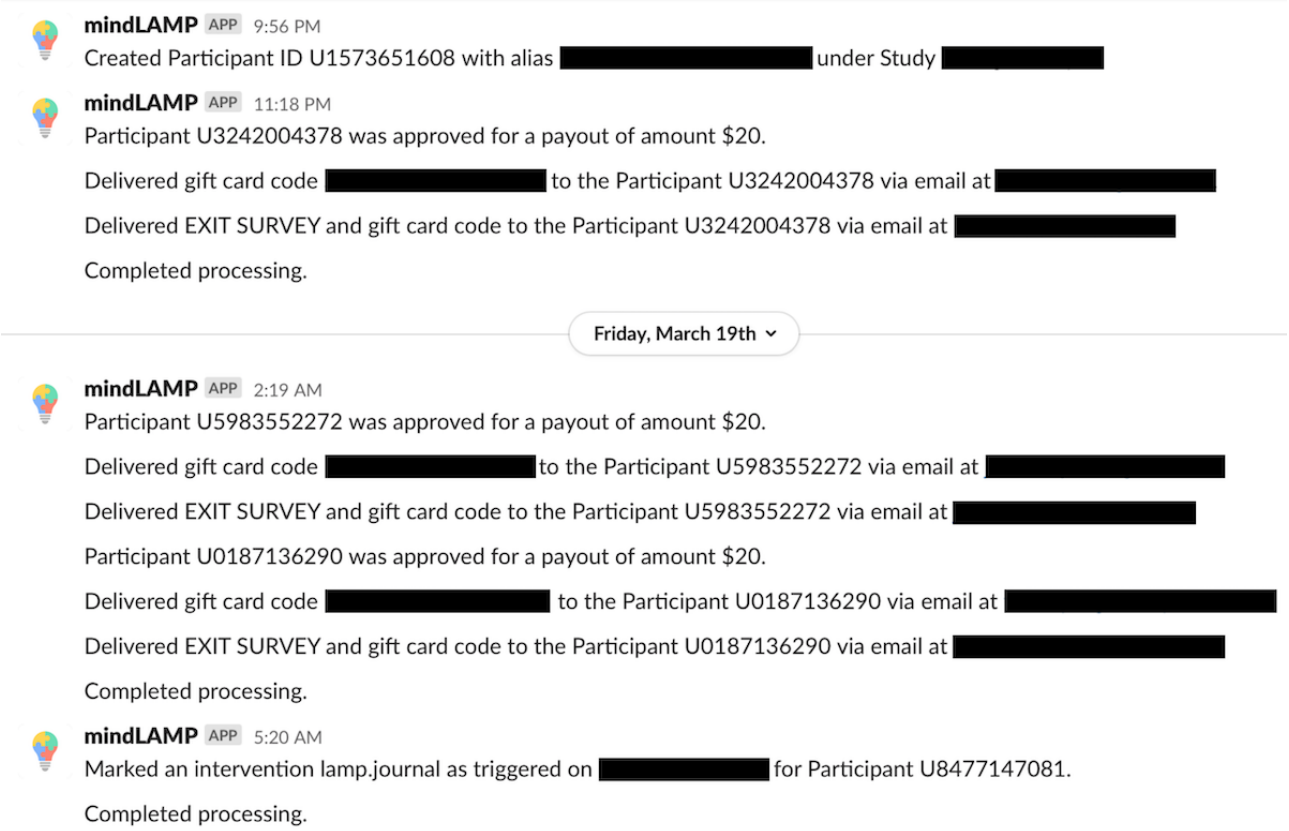


Figure 10. The detailed coordination required among the many components of the Learn, Assess, Manage, and Prevent platform involved in the submission of a push notification for a survey or gift card email for completion of a study; an example of the reporting of live intervention processing as made possible by a push-based model. Upon participant enrollment, survey delivery, gift card delivery, or intervention triggering, a message is pushed and logged to the Slack messaging service, a push-based model, to alert the research coordinator in real time. API: application programming interface; APNS: Apple Push Notification Service; FCM: Firebase Cloud Messaging; LAMP: Learn, Assess, Manage, and Prevent; REST: Representational State Transfer; SNS: Simple Notification Service.



XSL•FO
RenderX

Figure 11. An example of the reporting of live intervention processing as made possible by a push-based model. Upon subject enrollment, survey delivery, gift card delivery, or intervention triggering, a message is pushed and logged to the Slack messaging service, also a push-based model, to alert the research coordinator in real time.



The data collection processes are executed both actively during patient interactions with the mindLAMP app as well as passively while the app or mobile device is not in use. These data are securely uploaded to the organization's backend systems and can be used immediately upon receipt for data analysis and logic to select interventions to display to the patient. Using push notifications sent to the mobile device, the platform promotes a high level of engagement with patients without explicitly requiring approval for data upload. Once a research study or clinic is configured, its documenting configuration can be exported and reimported by other LAMP-compatible systems or interfaces. This allows reproducibility in both clinics and research studies, for example, by attaching the configuration file to a research manuscript or clinical protocol.

Consortium and Clinical Research Efforts

The LAMP platform is built and maintained collaboratively as an open platform to address the needs of many and integrate tools and resources to streamline workflows. In contrast to commercially available apps and services, the mindLAMP app may be used by organizations independently of our team through the deployment of a secure self-hosted backend. It can be customized and adapted without requiring specialized coding and deployment efforts, although others have also taken advantage of the extensibility of the platform to design and develop unique cognitive tests for their organization's needs. Common data processing and analysis needs across clinical and research workflows are encapsulated by the platform and the Cortex data analysis toolkit to minimize the time between patient onboarding and affecting or assessing patient outcomes.

The LAMP platform is highly configurable to suit many needs across a broad range of both clinical and research use cases and strategies. Consortium partners are encouraged to share their use case and LAMP configurations. As shown in Textbox 1, there were many different potential configurations and use patterns across consortium members.

XSL•FO
RenderX

Textbox 1. Selected examples of configurations and use cases for the Learn, Assess, Manage, and Prevent platform across various consortium members.

Ecological momentary assessment tool

- A total of 17 adults with substance use disorder who recently discharged from the hospital for this diagnosis completed daily assessments delivered via mindLAMP of mood, anxiety, sleep, social activity, and craving. No sensors were configured, and push notifications were enabled as reminders. Participants were able to view select survey responses in the Prevent tab. Results will be presented by the study team at the 34th Annual European College of Neuropsychopharmacology meeting in October 2021.

Digital phenotyping+

- A research study to examine circadian rhythms in bipolar disorder was conducted in which no activities were enabled or scheduled. The accelerometer, gyroscope, magnetometer, gravity, device motion, GPS, screen state, call and text, Bluetooth, and Wi-Fi sensors were enabled and configured to collect data at the highest possible frequency. Participants were able to view select sensor data in the Prevent tab.

Both ecological momentary assessment and digital phenotyping+

- A research study was conducted in which 100 college students were remotely enrolled to use the mindLAMP app for 1 month. Participants took 1 scheduled daily survey and 1 scheduled weekly survey, with provided optional tips and resources for managing stress, depression, and anxiety. The accelerometer, gyroscope, magnetometer, gravity, device motion, GPS, screen state, call and text, Bluetooth, and Wi-Fi sensors were enabled and configured to collect data at the highest possible frequency. Results are summarized in a paper published in 2021 [32].

Individual patient study

- A research study was conducted in which 50 participants with schizophrenia or bipolar disorder used the mindLAMP app for 1 year. Each participant was scheduled a daily standard battery of several surveys, the Jewels cognitive test, and the Spatial Span cognitive test. Optional tips and resources were provided, and the journaling, scratch card, and breathing exercise activities were made available for participants to use on their own volition. Participants were able to view their own data in the Prevent tab and worked with the research coordinator and psychiatrist to create custom surveys specific to individual participants' needs. For example, 1 participant chose to create a water intake survey. Each participant's notifications were scheduled individually by the research coordinator, instead of at a set time across all participants. The accelerometer, gyroscope, magnetometer, gravity, device motion, GPS, screen state, and call and text sensors were enabled and configured to collect data at the highest possible frequency.

Clinical use

- A clinical team in California offered dialectical behavioral therapy diary cards to all patients via mindLAMP. Before clinic visits and during their daily lives, patients would fill out these app-based dialectical behavioral therapy diary cards on their mobile device and were able to see their previous responses in the Prevent tab. No other activities were made available to the patients and no sensors were enabled for data collection. The diary cards were reviewed during each clinical session with the dialectical behavioral therapy therapist.

Digital clinic

- A clinic was established in which patients used the mindLAMP app with a new care team in addition to their ongoing care. Each patient was scheduled a daily standard battery of several surveys. Required tips and resources were provided along with required journaling, scratch card, and breathing exercise activities that were scheduled according to patient preferences. Patients were able to view their own data in the Prevent tab and worked with the digital navigator and clinician to create custom surveys or activities specific to individual patient's needs. For example, 1 patient requested a set of self-management tips and resources. Each patient's notifications were scheduled individually by the digital navigator. On the basis of each patient's clinical goals each week, accelerometer, gyroscope, magnetometer, gravity, device motion, GPS, screen state, call and text, and other sensors were turned off or on with the goal of capturing relevant and actionable information to help manage care. A protocol for the clinic is published here [33].

Intervention tool

- A research team used the cognitive games in mindLAMP as a tool for cognitive remediation to offer attention and memory training to patients with clinical high risk for psychosis. A paper summarizing the results was published in 2021 [34]. The app was offered in the Mandarin Chinese language for this study.

To aid these joint research and clinical efforts, the LAMP consortium was founded to connect partners using the LAMP platform. The design and development of the platform occurs in an open-source, collaborative environment that many have taken advantage of to suggest features, report bugs, add documentation, and improve the overall quality and efficacy of the LAMP platform. Through its flexibility and interoperability, the platform encourages integration and cross talk between clinical and research contexts, and to this end, supports the implementation of a digital clinic [33] and the creation of a digital navigator role [35].

Next Steps

The consortium further integrates directly into the development and feedback cycle using a community forum and bug tracking system, both available publicly. The community forum serves as a centralized resource for multiple teams or organizations to work with one another to assist with data analysis or troubleshooting and provide feedback about the LAMP platform. In addition, collaborators actively engage in making modifications to the source code (hosted through the public source code repository hosting service GitHub), make any suggested modifications or bug fixes, and then request that these

XSL•FO
RenderX

changes be merged *upstream* into the distribution of the LAMP platform that is used by all. To learn more about the LAMP platform or help contribute, one can join the consortium or visit the open-source repository [36].

Conclusions

Through the incorporation of patient- and clinician-centric feedback as well as a standards-based approach, the LAMP platform is designed to address important needs around the effective and compatible integration of PGHD into existing clinical systems for research and clinical care. It offers a flexible and comprehensive set of tools and solutions that can be configured and stitched together to function in a wide range of use cases, as used by members of the LAMP consortium. Its simplified programming interfaces are designed to securely handle a high throughput of PGHD as well as its companion metadata. With the integration of the Cortex data analysis toolkit, machine learning feature extraction, data processing, interactive visualization, and other essential tasks are simplified and coordinated seamlessly at low cost and high efficiency. In addressing technical challenges, the LAMP platform enables research and clinical teams to rapidly convert PGHD from widely accessible consumer smartphones and wearable devices into actionable clinical insights.

Acknowledgments

This work was supported by a philanthropic gift from Jeremy Wertheimer.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A full listing of Active, Passive, and Cortex data types currently supported by the Learn, Assess, Manage, and Prevent platform along with a description and expected components of the data.
[DOCX File , 18 KB - mhealth_v10i1e30557_app1.docx]

References

1. Patient-generated health data. HealthIT.gov. URL: https://www.healthit.gov/topic/scientific-initiatives/patient-generated-health-data [accessed 2021-12-15]
2. Chan S, Torous J, Hinton L, Yellowlees P. Mobile tele-mental health: increasing applications and a move to hybrid models of care. Healthcare (Basel) 2014 May 06;2(2):220-233 [FREE Full text] [doi: 10.3390/healthcare2020220] [Medline: 27429272]
3. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
4. Mobile fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-12-15]
5. About one-in-five Americans use a smart watch or fitness tracker. Pew Research Center. 2020. URL: https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ [accessed 2021-12-15]
6. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
7. Bardram JE, Matic A. A decade of ubiquitous computing research in mental health. IEEE Pervasive Comput 2020 Jan;19(1):62-72. [doi: 10.1109/mprv.2019.2925338]
8. Baxter C, Carroll J, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020 Feb 03;8(2):e16741 [FREE Full text] [doi: 10.2196/16741] [Medline: 32012102]
9. Torous J, Vaidyam A. Multiple uses of app instead of using multiple apps - a case for rethinking the digital health technology toolbox. Epidemiol Psychiatr Sci 2020 Jan 31;29:e100 [FREE Full text] [doi: 10.1017/S2045796020000013] [Medline: 32000876]
10. Demiris G, Iribarren SJ, Sward K, Lee S, Yang R. Patient generated health data use in clinical practice: a systematic review. Nurs Outlook 2019;67(4):311-330 [FREE Full text] [doi: 10.1016/j.outlook.2019.04.005] [Medline: 31277895]
11. Bender D, Sartipi K. HL7 FHIR: An Agile and RESTful approach to healthcare information exchange. In: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems. 2013 Presented at: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems; Jun 20-22, 2013; Porto, Portugal. [doi: 10.1109/cbms.2013.6627810]
12. ITI homepage. ITI. URL: https://www.itic.org/ [accessed 2021-12-15]

XSL•FO
RenderX

13. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]
14. Sayeed R, Gottlieb D, Mandl KD. SMART Markers: collecting patient-generated health data as a standardized property of health information technology. NPJ Digit Med 2020 Jan 23;3:9 [FREE Full text] [doi: 10.1038/s41746-020-0218-6] [Medline: 31993507]
15. United States Core Data for Interoperability (USCDI). HealthIT.gov. URL: https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi [accessed 2021-12-15]
16. mCODE™: Minimal Common Oncology Data Elements. mCODE. URL: https://mcodeinitiative.org/ [accessed 2021-12-15]
17. Dugas M, Neuhaus P, Meidt A, Doods J, Storck M, Bruland P, et al. Portal of medical data models: information infrastructure for medical research and healthcare. Database (Oxford) 2016 Feb 11;2016:bav121 [FREE Full text] [doi: 10.1093/database/bav121] [Medline: 26868052]
18. Lack of EHR interoperability, API use hinder value-based care. EHR Intelligence. URL: https://ehrintelligence.com/news/lack-of-ehr-interoperability-api-use-hinder-value-based-care [accessed 2021-12-15]
19. Torous J, Wisniewski H, Bird B, Carpenter E, David G, Elejalde E, et al. Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach. J Technol Behav Sci 2019 Apr 27;4(2):73-85 [FREE Full text] [doi: 10.1007/s41347-019-00095-w]
20. Rodriguez-Villa E, Mehta UM, Naslund J, Tugnawat D, Gupta S, Thirtalli J, et al. Smartphone Health Assessment for Relapse Prevention (SHARP): a digital solution toward global mental health. BJPsych Open 2021 Jan 07;7(1):e29 [FREE Full text] [doi: 10.1192/bjo.2020.142] [Medline: 33407986]
21. Hoffman L, Wisniewski H, Hays R, Henson P, Vaidyam A, Hendel V, et al. Digital Opportunities for Outcomes in Recovery Services (DOORS): a pragmatic hands-on group approach toward increasing digital health and smartphone competencies, autonomy, relatedness, and alliance for those with serious mental illness. J Psychiatr Pract 2020 Mar;26(2):80-88 [FREE Full text] [doi: 10.1097/PRA.0000000000000450] [Medline: 32134881]
22. Rauseo-Ricupero N, Henson P, Agate-Mays M, Torous J. Case studies from the digital clinic: integrating digital phenotyping and clinical practice into today's world. Int Rev Psychiatry 2021 Jun;33(4):394-403. [doi: 10.1080/09540261.2020.1859465] [Medline: 33792463]
23. OpenAPI specification. Swagger. URL: https://swagger.io/specification/ [accessed 2021-12-15]
24. JSON Schema: a media type for describing JSON documents. JSON Schema. URL: https://json-schema.org/draft/2020-12/json-schema-core.html [accessed 2021-12-15]
25. Representational State Transfer (REST). University of California. URL: https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm [accessed 2021-12-15]
26. The JavaScript Object Notation (JSON) data interchange format. Internet Engineering Task Force. URL: https://datatracker.ietf.org/doc/html/rfc7159 [accessed 2021-12-15]
27. The Transport Layer Security (TLS) protocol version 1.3. Internet Engineering Task Force. URL: https://datatracker.ietf.org/doc/html/rfc8446 [accessed 2021-12-15]
28. Advanced Encryption Standard (AES). Information Technology Laboratory. URL: https://csrc.nist.gov/publications/detail/fips/197/final [accessed 2021-12-15]
29. HTML Living Standard. HTML. URL: https://html.spec.whatwg.org [accessed 2021-12-15]
30. Vega. URL: https://vega.github.io [accessed 2021-12-15]
31. What is Kubernetes? Docker. URL: https://www.docker.com/products/kubernetes [accessed 2021-12-15]
32. Melcher J, Lavoie J, Hays R, D'Mello R, Rauseo-Ricupero N, Camacho E, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. J Am Coll Health 2021 Mar 26:1-13. [doi: 10.1080/07448481.2021.1905650] [Medline: 33769927]
33. Rodriguez-Villa E, Rauseo-Ricupero N, Camacho E, Wisniewski H, Keshavan M, Torous J. The digital clinic: implementing technology and augmenting care for mental health. Gen Hosp Psychiatry 2020;66:59-66 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.06.009] [Medline: 32688094]
34. Li H, Yang S, Chi H, Xu L, Zhang T, Singleton G, et al. Enhancing attention and memory of individuals at clinical high risk for psychosis with mHealth technology. Asian J Psychiatr 2021 Apr;58:102587. [doi: 10.1016/j.ajp.2021.102587] [Medline: 33618070]
35. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020 Nov 26;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
36. LAMP Platform. Zenodo. URL: https://zenodo.org/record/4716576#.YboR4L1BzIU [accessed 2021-12-15]

Abbreviations

API: application programming interface
FHIR: Fast Healthcare Interoperability Resources

XSL•FO
RenderX

IDE: integrated development environment
LAMP: Learn, Assess, Manage, and Prevent
PGHD: patient-generated health data
SMART: Specific, Measurable, Achievable, Realistic, and Timely

Edited by L Buis; submitted 19.05.21; peer-reviewed by J Oldenburg, B Chaudhry; comments to author 13.08.21; revised version received 18.08.21; accepted 11.11.21; published 07.01.22.

Please cite as:
Vaidyam A, Halamka J, Torous J
Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study
JMIR Mhealth Uhealth 2022;10(1):e30557
URL: https://mhealth.jmir.org/2022/1/e30557
doi:10.2196/30557
PMID:34994710

©Aditya Vaidyam, John Halamka, John Torous. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 07.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Daily Level Association of Physical Activity and Performance on Ecological Momentary Cognitive Tests in Free-living Environments: A Mobile Health Observational Study

Zvinka Z Zlatar[1], PhD; Laura M Campbell[2], MS; Bin Tang[1], PhD; Spenser Gabin[3], MA; Anne Heaton[1], BA; Michael Higgins[4], MS; Joel Swendsen[5,6], PhD; David J Moore[1], PhD; Raeanne C Moore[1], PhD

[1]Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States
[2]San Diego Joint Doctoral Program in Clinical Psychology, San Diego State University/University of California, San Diego, CA, United States
[3]Department of Counseling and Marital and Family Therapy, University of San Diego, San Diego, CA, United States
[4]Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego, La Jolla, CA, United States
[5]National Center for Scientific Research, University of Bordeaux, Bordeaux, France
[6]Ecole Pratique des Hautes Etudes PSL Research University, Paris, France

Corresponding Author:
Zvinka Z Zlatar, PhD
Department of Psychiatry
University of California, San Diego
9500 Gilman Dr
MC 0811
La Jolla, CA, 92093
United States
Phone: 1 858 822 7737
Email: zzlatar@health.ucsd.edu

Abstract

Background: Research suggests that physical activity (PA) has both acute and chronic beneficial effects on cognitive function in laboratory settings and under supervised conditions. Mobile health technologies make it possible to reliably measure PA and cognition in free-living environments, thus increasing generalizability and reach. Research is needed to determine whether the benefits of PA on cognitive function extend from the laboratory to real-world contexts.

Objective: This observational study aims to examine the association between daily fluctuations in PA and cognitive performance using mobile health technologies in free-living environments.

Methods: A total of 90 adults (mean age 59, SD 6.3 years; 65/90, 72% men) with various comorbidities (eg, cardiovascular risk and HIV) and different levels of baseline cognition (ranging from cognitively normal to impaired) completed ecological momentary cognitive tests (EMCTs) on a smartphone twice daily while wearing an accelerometer to capture PA levels for 14 days. Linear mixed-effects models examined the daily associations of PA with executive function and verbal learning EMCTs. Moderation analyses investigated whether the relationship between daily PA and daily performance on EMCTs changed as a function of baseline cognition, cardiovascular risk, and functional status (independent vs dependent).

Results: Days with greater PA were associated with better (faster) performance on an executive function EMCT after covariate adjustment (estimate −0.013; β=−.16; *P*=.04). Moderation analyses (estimate 0.048; β=.58; *P*=.001) indicated that days with greater PA were associated with better (faster) executive function performance in individuals who were functionally dependent (effect size −0.53; *P*<.001) and not in functionally independent adults (effect size −0.01; *P*=.91).

Conclusions: EMCTs may be a sensitive tool for capturing daily-level PA-related fluctuations in cognitive performance in real-world contexts and could be a promising candidate for tracking cognitive performance in digital health interventions aimed at increasing PA. Further research is needed to determine individual characteristics that may moderate the association between daily PA and EMCT performance in free-living environments.

(JMIR Mhealth Uhealth 2022;10(1):e33747) doi:10.2196/33747

XSL•FO
RenderX

KEYWORDS

smartphones; neuropsychology; ecological momentary assessment; digital health; exercise; people living with HIV; aging; wearables; mobile cognition; mobile phone

Introduction

Background

The exponential growth of the older adult population will result in more individuals living with Alzheimer's disease and related dementias (ADRD) [1]. Given the lack of effective treatments for Alzheimer's disease, a focus on healthy lifestyle choices holds promise in preventing cognitive decline [2-4]. Evidence links physical activity (PA) with long-term cognitive benefits such as reduction of age-related cognitive decline and dementia risk. For example, a meta-analysis of longitudinal observational studies with middle-aged and older adults found that higher levels of PA were associated with a 14% lower risk of dementia when compared to those with lower levels of PA [5]. Similarly, compared with sedentary individuals, even low to moderate PA has been associated with a 35% reduction in the risk of cognitive decline [6]. Moreover, a recent review [7] found a small positive effect of PA on executive function (Cohen's d=0.27) and memory (Cohen's d=0.24), whereas another study found that walking more steps per day was related to better executive function in healthy aging [8]. Several mechanisms explain how PA benefits cognition in the long term, including angiogenesis, neurogenesis, upregulation of neurotrophic factors, reduced inflammation, cardiovascular benefits, changes in central arterial stiffness and endothelial function, and insulin regulation [9,10].

Less is known about the acute effects of PA on cognition in middle-aged and older adults in real-world contexts, where ecological validity is optimized. The acute effects of PA on cognition have generally been studied in laboratory settings, where individuals complete cognitive tasks before, during, or immediately following an exercise challenge. Within these settings, it has been found that acute moderate intensity PA results in improved executive function [11-13] and working memory performance in older adults [14], although others have found that acute exercise before memory encoding tasks may impair performance in older adults [15]. A meta-analysis of the literature showed that the overall effect of acute exercise on cognition is positive but generally small; and that longer exercise duration, greater intensity, type of cognitive performance assessed (and when it is assessed), and greater fitness level appear to be significant moderators of larger effect sizes [16].

Owing to research-grade accelerometry, it is now possible to remotely track PA behavior, whereas smartphone-based technology can assess cognitive function in real-world contexts using ecological momentary cognitive tests (EMCTs). Evidence suggests that EMCTs have relatively high completion rates of 60% to 85% [17], and performance on these tests correlates with standard neuropsychological testing scores [18-22]. With 81% of adults aged 60 to 69 years and 62% of those aged ≥70 years owning a smartphone [23,24], EMCTs can be deployed to increase our understanding of whether acute fluctuations in real-world PA lead to immediate improvements in cognition in a variety of everyday contexts. This can help determine the utility of EMCTs as tools in digital health interventions to measure cognitive function repeatedly, at scale, and in people's natural environments, thus improving ecological validity and reducing participant burden.

Objective

This observational study investigates if there are daily-level associations between accelerometer-measured PA and smartphone-based EMCT performance during a 2-week measurement period in middle-aged and older adults with a variety of comorbid conditions (eg, people living with HIV [PWH], cardiovascular risk, and functional impairment). Moreover, we examine whether cognitive status, functional status, and cardiovascular disease (CVD) risk affect the acute associations of PA with performance on mobile verbal learning and executive function EMCTs. We hypothesize that days with greater measured PA would be associated to better performance on executive function and learning. Consistent with the literature suggesting that those with greater risk profiles may benefit most from lifestyle approaches to maintain brain health [25], we also hypothesize that the association between daily fluctuations in PA and EMCT performance would be stronger in participants with lower cognition, higher CVD risk, and lower functional independence status compared to those with better cognition, lower CVD risk, and no functional dependence.

Methods

Participants

A total of 90 people—57 (63%) PWH and 33 (37%) HIV-negative, middle-aged, and older adults aged 50 to 74 years—were enrolled in a study at the University of California San Diego's HIV Neurobehavioral Research Program (HNRP) from 2016 to 2019.

Recruitment

Participants were recruited from the participant pool at the HNRP or through community outreach. Inclusion criteria were age ≥50 years, fluent in English, and ability to provide written informed consent. Exclusion criteria were neurological confounders not related to HIV (eg, stroke, untreated seizure disorder, and head injury with loss of consciousness >30 minutes), diagnosis of serious mental illness (eg, schizophrenia and bipolar disorder), and a reported learning disability or low estimated verbal IQ (ie, a standard score <70 on the Wide Range Achievement Test 4 Reading test [26]). The overall level of illicit substance use in this sample was low [27]. Illicit substance use was not exclusionary; however, participants with a positive alcohol breathalyzer or positive urine toxicology (other than marijuana or prescription medications) on the day of testing were rescheduled for another day. This occurred 2 times, and these 2 participants were rescheduled and tested positive once more upon their return, at which time they were withdrawn from the study. The study procedures were approved by the institutional review board of the University of California, San

XSL•FO
RenderX

Diego, and all participants demonstrated decisional capacity [28] and provided written informed consent.

Measures and Procedures

Study Overview

The study comprised a baseline in-person visit, a 14-day period of smartphone-based EMCTs and wrist-worn accelerometer tracking in participants' natural environments, and a follow-up in-person visit. Participants were not co-enrolled in other research during the study period. Participants were compensated for the in-person assessments, for each EMCT that they completed, and for returning the study-owned smartphones and accelerometer watches. During the baseline visit, participants were given a wrist-worn accelerometer to track PA and a smartphone (Samsung Galaxy S 4.2 YP-GII) on which daily cognitive tests were administered. They received a 20- to 30-minute tutorial with an examiner on the use of the smartphone and how to complete the EMCTs. To ensure security of the data, the study phone's operating system was encrypted in case the smartphone was lost or stolen. A user manual was sent home with participants, which included information on using the smartphone, proper wear of the accelerometer, frequently asked questions, and study staff contact information.

Baseline Laboratory Visit

Evaluation of Baseline Cognition

At the baseline visit, all participants completed a standardized and comprehensive neuropsychological battery used in research studies at the HNRP [29] covering 7 cognitive domains, including verbal fluency, speed of information processing, learning, delayed recall, working memory, executive function, and motor skills (Multimedia Appendix 1). Raw scores on these cognitive assessments were converted to practice effect–corrected scaled scores (mean 10, SD 3) to control for prior exposure to neuropsychological testing [30]. Next, scaled scores were converted to demographically adjusted (age, sex, education, and race) T scores for each test (T score <40 indicates impaired performance). Adjusted T scores were averaged to compute the global T score [31,32], which was the variable of interest in later analyses.

Evaluation of Functional Status

Participants completed the *Lawton–Brody instrumental activities of daily living* (IADL) questionnaire [33] to assess their ability to function independently in everyday life. The measure requires the participants to rate their ability to complete various basic (eg, bathing and dressing) and instrumental (eg, managing finances and cooking) activities of daily living both at their current level and *best ever* level of functioning. A decline in functioning is indicated if the current level of functioning is lower than that of the *best level* on any task. A participant with ≥2 declines is deemed *IADL-dependent*.

Evaluation of Cardiovascular Risk

To measure cardiovascular risk, we calculated the participants' *Framingham CVD risk* scores [34]. This CVD risk score assigns weighted point values to the following factors: age, diabetes status, smoking status, systolic blood pressure (treated vs untreated), total cholesterol, and high-density lipoprotein cholesterol level to create a risk estimate score. The equation for calculating the Framingham CVD risk score is provided in the study by D'agostino et al [34].

Neuromedical Evaluation

During the baseline laboratory visit, participants completed a neuromedical evaluation, which included a fasting blood draw, in which an HIV and hepatitis C virus antibody point-of-care rapid test (Miriad-MedMira) was conducted and confirmed with western blot analyses. Blood samples were used to measure the current CD4 T-cell counts and plasma HIV viral loads (detectable at >50 copies/mL). All participants completed a neuromedical interview, including the collection of medical history, medication lists, and other HIV disease characteristics (ie, AIDS status, estimated duration of HIV disease, nadir CD4 T-cell count, and history of antiretroviral therapy).

At-home Monitoring: 14-Day Mobile Assessment

PA Measurement

To measure objective PA, participants wore the ActiGraph GT9X Link device (ActiGraph Inc) continuously on their nondominant wrist, except while bathing or swimming, for the duration of the 14-day EMCT period. Participants were also asked to record when and why they took the device off. The ActiGraph Link (ActiGraph Inc) is a triaxial accelerometer that has consistently been shown to be a valid and reliable measure of PA [35-37]. The wear location and assessment period are aligned with the best practices for PA assessment, resulting in high levels of acceptability and compliance among participants [38-41]. The device is small and lightweight and has a minimum of 512 MB of nonvolatile flash memory. The data are stored on the device. When participants returned the device, the ActiGraph data were immediately downloaded and screened by hour for completeness and possible irregularities or malfunction according to best practice recommendations [37,39,40]. Participants' data were included if they wore the device for a minimum of 5 days, with at least 600 minutes of wear each day. Participants who achieved 4 days of wear, with at least 3000 total minutes, were also included. PA was defined as cumulative activity *counts* (ActiGraph's proprietary metric) per day. This metric incorporates intensity, frequency, and duration of acceleration and is recommended for assessing the total volume of PA in a 24-hour period [42]. Specifically, counts are a result of summing the postfiltered accelerometer values (raw data at 30 Hz) into epoch *chunks*. The value of the counts varies based on the frequency and intensity of the raw acceleration. Vector magnitude was then calculated using the following equation:

Counts per minute (CPM) is the average amount of total movement throughout the day; the higher the CPM, the more activity throughout the day. CPM is used as there are no established cutoff points for measuring sedentary, light, or moderate-to-vigorous PA at the wrist. Assessing the total volume of PA via vector magnitude CPM is important as it takes the frequency, intensity, and duration of activity bouts and condenses them down to a single metric that can be harmonized across studies.

XSL•FO
RenderX

EMCT Paradigm

An alarm sounded on the study smartphone twice a day for the 14-day assessment period to signal when it was time to complete 1 of the 2 mobile cognitive tests: once for the mobile color–word interference test (mCWIT) [18] and once for the mobile verbal learning test (mVLT; Figure 1) [19]. The alarms occurred at pseudorandom times throughout the day, accounting for participants' preferred sleep-wake schedules, such that the participant did not know when they would be asked to complete the tests. Alarms occurred 2 to 3 hours apart. Once the alarm sounded, participants had 10 minutes to start the assessment (with a reminder alarm every 2 minutes during that 10-minute window) before it would time-out and be considered missed. Participants also had the option to cancel the survey during the 10-minute window or at any point during the survey. The mCWIT and mVLT were never given at the same time point.

The mCWIT is a test based on the Stroop paradigm assessing executive function. A total of 16 words (4 rows of 4 words) are presented in a different color than the printed word, and participants are instructed, "Do not read the words, say the colors in which they are written." There was one trial for this test at each administration, for which participants had up to 60 seconds to say the colors for all 16 items as fast as possible. Each administration of the mCWIT alternated the order of words as well as the nonmatching color of each word. Responses were audio recorded and scored by 2 independent raters. All discrepant scores were assessed by a third rater. The outcome assessed in this study was the completion time (seconds).

Figure 1. Example screenshot of the mobile color–word interference test (mCWIT; test of executive function; left panel) and the mobile verbal learning test (mVLT; test of verbal learning; right panel). The words on the mVLT are sample words and not an actual word list from the mVLT. All 12 words are presented on the screen, eliminating a need to scroll to view all the words.

The mVLT is a test designed to assess verbal learning and recall. Participants were presented with a list of 12 semantically unrelated words to read on the smartphone for three 30-second learning trials during which participants were told to memorize the words. After each of the 3 learning trials, participants were asked to immediately recall as many words as possible by saying them out loud into the phone. Participants had 60 seconds to recall words before the next trial began. A unique list was presented each day. Responses were audio recorded and scored for the total number of correct responses. The outcome assessed in this study was the total number of words correctly (out of a possible 36 words) recalled over the 3 learning trials per day. The mVLT was scored by 2 independent raters, and discrepant scores were reviewed by an additional rater.

For both the mCWIT and the mVLT, trials were excluded from analyses if raters suspected cheating (eg, help from others) or if the participant was doing something else during the test (eg, talking with others). mCWIT data were excluded from analyses for 2% (2/90) of participants; one of the participants was excluded because of colorblindness, and another participant was excluded as they had an average of 15/16 errors, and therefore, their data were not considered valid. There were also 2% (2/90) of participants who had 14% mCWIT compliance (the minimal compliance threshold of 30% was set for analyses) and were thus removed from the analyses.

Statistical Analysis

Demographic data, comorbid conditions, and HIV disease characteristics were summarized as mean (SD), median (IQR), or n (%). mCWIT in seconds was $\log_{10}$-transformed to improve normality before the analyses. To evaluate whether days with greater PA were associated with better EMCT performance, linear mixed-effects models with subject-specific random intercepts were conducted to detect the relationship between within-person PA and EMCT performance. Between-person PA effects on EMCT performance were also estimated. Crude models with significant within- or between-person effects ($P<.05$) were later adjusted for covariates. HIV status, age, sex, education, and race or ethnicity (non-Hispanic White vs all other race or ethnicities) were included as covariates in the models if $P<.10$, using backward model selection.

In addition, to determine whether cognition (ie, global T score), functional status (ie, IADL dependent or independent), and CVD risk (Framingham CVD risk profile) moderated the relationship between PA and EMCT performance, moderation analyses were conducted using separate linear mixed-effects models with fixed effects of within- and between-person PA,

XSL•FO
RenderX

global T score, and IADL status or Framingham stroke risk score; their interactions (eg, an interaction between within-person PA and global T score on EMCT performance); and subject-specific random effects. All models were controlled for study day. The significance level of α was set at .05. The effect sizes (and 95% CIs) are standardized coefficients, which are analogous to Cohen's *d*. R software (version 3.6.0; R Foundation for Statistical Computing) was used to perform all statistical analyses.

Results

Participants

Participants' demographic, clinical, and cognitive characteristics are presented in Table 1. On average, participants were in their late 50s (mean age 59, SD 6.3 years), had some college education (mean 14.4, SD 2.4 years), and most were male (65/90, 72%) and identified as non-Hispanic White (57/90, 63%). Of the 90 participants, 57 (63%) were PWH, and participants' age, education, and race or ethnicity were not significantly different based on HIV status. There were significantly more women in the HIV-negative group compared with the PWH group (PWH: 10/57, 18% women; HIV-negative: 15/33, 45% women; $P<.004$). Overall, there was excellent adherence to the mobile cognitive testing and accelerometer protocols as, on average, participants completed 13 of 14 days of the mCWIT, 12 of 14 days of the mVLT, and had 12 of 14 days of accelerometer data to objectively assess PA.

XSL•FO
RenderX

Table 1. Participant characteristics (N=90).

Demographic variables	Values
Age (years), mean (SD)	59.0 (6.3)
Male, n (%)	65 (72)
Race and ethnicity, n (%)	
Non-Hispanic White	57 (63)
African American or Black	19 (21)
Hispanic or Latino	11 (12)
Other	3 (3)
Education (years), mean (SD)	14.4 (2.4)
Comorbid conditions	
Hyperlipidemia, n (%)	52 (58)
Hypertension, n (%)	50 (56)
Diabetes mellitus, n (%)	21 (23)
BMI (n=80), median (IQR)	27.0 (25.1-32.5)
Framingham cardiovascular disease risk score (n=80), median (IQR)	16.1 (9.4-27.33)
Depression (Beck Depression Inventory-2), median (IQR)	4 (1-10)
Lifetime any substance use disorder, n (%)	57 (63)
Current substance use disorder[a], n (%)	3 (3)
HIV characteristics[b]	
AIDS, n (%)	39 (68)
Current CD4, median (IQR)	690 (549-879)
Nadir CD4, median (IQR)	145 (35-300)
Duration of HIV infection (years), median (IQR)	25.2 (18.5-28.9)
On antiretroviral therapy, n (%)	53 (93)
Undetectable viral load (n=53), n (%)	52 (98)
Cognitive variables	
Global T score, mean (SD)	49.5 (6.2)
Independent activities of daily living: dependent, n (%)	30 (33)
Mobile color–word interference test score (seconds; n=88), mean (SD)	23.4 (6.1)
Number of mobile color–word interference test trials completed (n=88), median (IQR)	13 (11-13.25)
Mobile verbal learning test score (total correct), mean (SD)	18.9 (4.8)
Number of mobile verbal learning test trials completed, median (IQR)	12 (11-13.25)
Physical activity	
Average vector magnitude counts per minute/day, mean (SD)	2047 (616.7)
Wear time (days), median (IQR)	12 (11-14)

[a]All current substance use disorders were cannabis or alcohol use disorders.

[b]Approximately 63% (57/90) were HIV-positive.

Within- and Between-Person Associations of PA and EMCT Performance

Overview

Results and effect sizes are presented in Table 2.

XSL•FO
RenderX

Table 2. Mixed-effects models for associations between physical activity and ecological momentary cognitive testing performance.

Models	Estimate (95% CI)	Effect size[a] (95% CI)	*P* value
Mobile color–word interference test of executive function (log_{10} transformed)			
Model 1 crude			
Within-person physical activity[b]	−0.013 (−0.0025 to −0.00005)	−0.15 (−0.30 to 0.0004)	.049
Between-person physical activity	−0.013 (−0.048 to 0.023)	−0.15 (−0.59 to 0.29)	.50
Model 1 adjusted[c]			
Within-person physical activity	−0.013 (−0.026 to −0.0008)	−0.16 (−0.31 to −0.008)	.04
Between-person physical activity	−0.004 (−0.038 to 0.030)	−0.049 (−0.48 to 0.38)	.82
Mobile verbal learning test			
Model 2 crude[d]			
Within-person physical activity	−0.52 (−1.37 to 0.33)	−0.093 (−0.25 to 0.061)	.24
Between-person physical activity	0.051 (−1.53 to 1.64)	0.009 (−0.28 to 0.29)	.95

[a]Standardized regression coefficient analogous to Cohen's *d*.

[b]Within-person and between-person physical activity is reflected as counts per minute/1000. All analyses control for study day.

[c]Adjusted for HIV status, age (in years), and race or ethnicity (reference: non-Hispanic White); the interaction between HIV status and physical activity was not significant and was therefore removed from the adjusted model.

[d]Model 2 was not subsequently adjusted for covariates because of the crude model being nonsignificant.

mCWIT Test of Executive Function Model 1 Crude

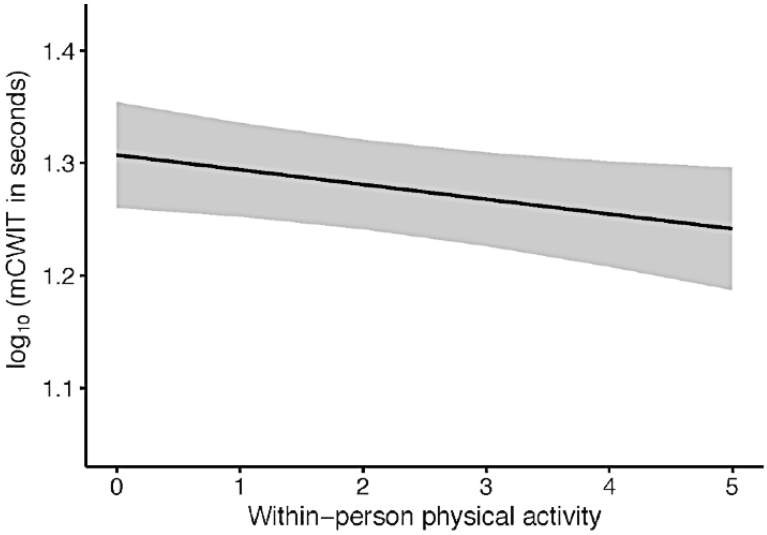
Within-person PA was significantly associated with mCWIT performance such that greater daily PA was associated with faster (ie, better) daily mCWIT performance. The between-person association of PA and mCWIT was not significant, indicating that average PA was not significantly associated with average mCWIT performance in this sample.

mCWIT Test of Executive Function Model 1 Adjusted

The significant within-person effect held after accounting for HIV status and the demographic variables that significantly improved the model (ie, age and race or ethnicity). The results indicate that for every increase of 1 SD in PA, mCWIT performance (on the log_{10} scale) improved (was faster) by 0.16 SDs, or mCWIT performance in seconds decreased (was faster) by 3% for every 1000-unit increase in PA (Figure 2). As with the crude model, the between-person association of PA and mCWIT was not significant.

To determine whether the significant within-person effects varied by HIV status, we examined the interaction term of HIV status by within-person PA. The interaction term did not reach significance (estimate 0.002; *P*=.88; effect size −0.025), indicating that the within-person association between daily PA and daily mCWIT performance did not differ by HIV status. Therefore, the HIV by PA interaction was not included in subsequent models examining the mCWIT.

Figure 2. Greater daily physical activity was associated with faster daily mobile color-word interference test of executive function performance. Within-person physical activity is reflected as counts per minute/1000. Variables were adjusted for between-person physical activity, HIV status, age, ethnicity, and study day. Shaded bands represent 95% CIs. mCWIT: mobile color-word interference test.



XSL•FO
RenderX

mVLT Test of Verbal Learning Model 2 Crude

Neither the within nor the between-person effects of PA with mVLT performance were significant, indicating that neither daily nor average PA was related to mVLT performance.

Sensitivity Analyses: Global Cognition, Functional Status, and Cardiovascular Risk

Overview

The results of the sensitivity analyses are presented in Table 3. We examined whether global cognition (ie, global T score), functional status (ie, IADL status), and CVD risk (ie, Framingham CVD risk score) moderated the relationship of within- or between-person PA and mobile cognitive test performance. Only models with significant interaction terms were adjusted for covariates.

XSL•FO
RenderX

Table 3. Mixed-effects models to examine whether cognition (global T score), Lawton–Brody instrumental activities of daily living (IADL), and cardiovascular risk (Framingham cardiovascular disease [CVD] risk score) moderated the association of physical activity (PA) and ecological momentary cognitive testing performance.

Model	Estimate (95% CI)	Effect size[a] (95% CI)	*P* value
mCWIT[b] test of executive function (log_{10} transformed)			
Model 3: global T score			
Global T score	−0.002 (−0.015 to 0.010)	−0.030 (−0.18 to 0.12)	.70
Within-person PA[c]×global T score	−.0007 (−0.003 to 0.001)	−0.008 (−0.033 to 0.017)	.52
Between-person PA×global T score	−.003 (−0.008 to 0.003)	−0.031 (−0.100 to 0.039)	.39
Model 4 crude: IADL[d]			
IADL status (reference: dependent)	0.046 (−0.12 to 0.21)	0.57 (−1.45 to 2.59)	.58
Within-person PA×IADL status	0.049 (0.021 to 0.077)	0.60 (0.26 to 0.95)	.001
Between-person PA×IADL status	−0.026 (−0.10 to 0.049)	−0.32 (−1.25 to 0.62)	.51
Model 4 adjusted[e]: IADL			
IADL status (reference: dependent)	0.018 (−0.14 to 0.17)	0.22 (−1.68 to 2.12)	.82
Within-person PA IADL status	0.048 (0.019 to 0.075)	0.58 (0.24 to 0.92)	.001
Between-person PA×IADL status	−0.005 (−0.074 to 0.064)	−0.065 (−0.94 to 0.81)	.89
Model 5: Framingham CVD risk score			
CVD risk	0.003 (−0.004 to 0.011)	0.038 (−0.054 to 0.13)	.42
Within-person PA×CVD risk	0.0002 (−0.0008 to 0.001)	0.002 (−0.009 to 0.014)	.70
Between-person PA×CVD risk	−0.001 (−0.005 to 0.002)	−0.016 (−0.062 to 0.030)	.50
mVLT[f] test of verbal learning			
Model 6: global T score			
Global T score	0.49 (−0.078 to 1.06)	0.088 (−0.016 to 0.19)	.10
Within-person PA×global T score	−0.042 (−0.18 to 0.090)	−0.008 (−0.032 to 0.016)	.54
Between-person PA×global T score	−0.13 (−0.39 to 0.13)	−0.023 (−0.070 to 0.025)	.35
Model 7: IADL			
IADL status (reference: dependent)	5.59 (−1.49 to 12.7)	1.00 (−0.29 to 2.29)	.13
Within-person PA×IADL status	0.67 (−1.30 to 2.63)	0.12 (−0.23 to 0.47)	.50
Between-person PA×IADL status	−2.60 (−5.90 to 0.69)	−0.47 (−1.07 to 0.13)	.13
Model 8: Framingham CVD risk score			
CVD risk	0.068 (−0.26 to 0.40)	0.012 (−0.048 to 0.073)	.69
Within-person PA×CVD risk	0.009 (−0.058 to 0.079)	0.002 (−0.011 to 0.014)	.79
Between-person PA×CVD risk	−0.052 (−0.22 to 0.11)	−0.010 (−0.040 to 0.021)	.54

[a]Analogous to Cohen's *d* (standardized regression coefficient).

[b]mCWIT: mobile color–word interference test.

[c]Within-person and between-person PA was reflected as counts per minute/1000. All analyses control for study day.

[d]Lawton–Brody instrumental activities of daily living questionnaire.

[e]Adjusted for HIV status, age (in years), and race or ethnicity (reference: non-Hispanic White).

[f]mVLT: mobile verbal learning test.

Relationships With Cognitive Global T Score

When predicting the mCWIT, the within-person PA by cognitive global T score and the between-person PA by cognitive global T score interactions were not significant (Table 3). When removing interaction terms, within-person PA remained significantly associated with mCWIT performance after accounting for cognitive global T score (estimate −0.013; *P*=.04). Examining the mVLT, neither the within-person PA

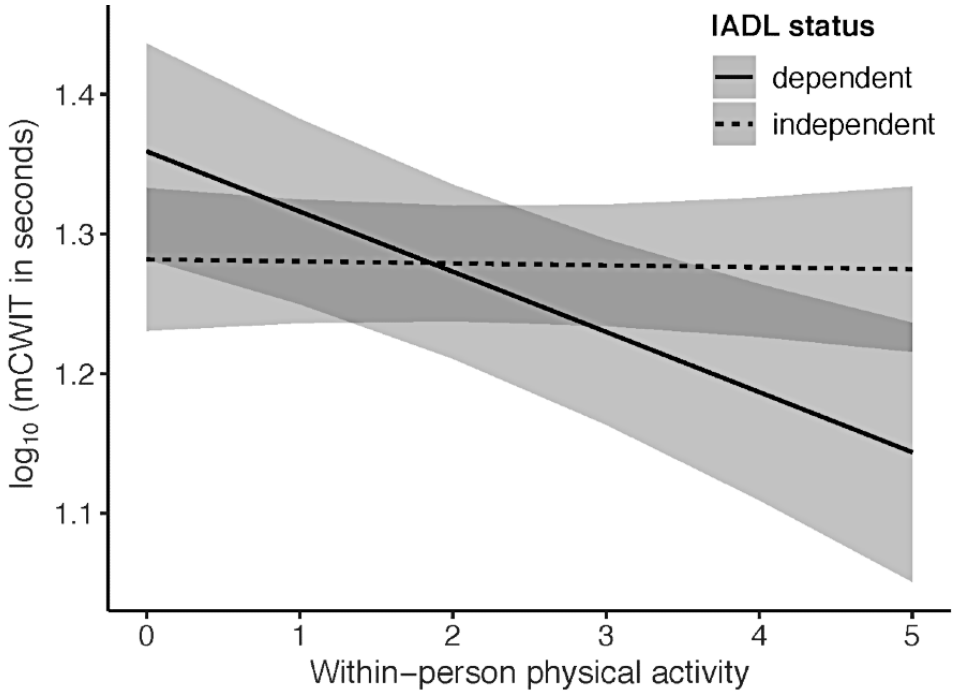
XSL•FO
RenderX

by cognitive global T score nor the between-person PA by cognitive global T score interactions were significant.

Relationships With IADL Status

There was a significant within-person PA by IADL score interaction effect on mCWIT performance. The relationship between greater daily PA and better mCWIT performance was significant for those who reported IADL dependence (effect size −0.53; *P*<.001) and not for those who were independent in IADL (effect size −0.01; *P*=.91; Figure 3). For those who reported IADL dependence, mCWIT performance decreased on average by 9.6% for every 1000 units of within-person PA. The interaction term remained significant when adjusting for demographic variables that significantly improved model fit (ie, age and race) and HIV status. The between-person PA by IADL score interaction was not significantly associated with mCWIT performance. Neither the between-person nor the within-person PA by IADL score interactions were significantly associated with mVLT performance.

Figure 3. Instrumental activities of daily living moderate the relationship of within-person physical activity and mobile color–word interference test of executive function performance. Within-person physical activity is reflected as counts per minute/1000. Variables were adjusted for between-person physical activity, HIV status, age, ethnicity, and study day. Shaded bands represent 95% CIs. IADL: instrumental activities of daily living; mCWIT: mobile color-word interference test.



Relationships With Framingham CVD Risk Score

For the mCWIT, the cardiovascular risk by within- and between-person PA interactions were not significant. After removing the interactions, within-person PA remained significantly related to mCWIT performance, even when accounting for cardiovascular risk (estimate −0.014; *P*=.048). When examining the mVLT, neither the within- nor between-person PA by cardiovascular risk interactions was significant.

Discussion

Principal Findings

We examined the cross-sectional association of objectively measured daily PA with EMCT performance to determine whether natural variation in PA was accompanied by fluctuations in performance on EMCTs of executive function and verbal learning in individuals' free-living environments. In a sample of diverse adults with a wide range of comorbidities (ie, cardiovascular risk and HIV), we found that during days with greater PA (ie, a combination of higher frequency, intensity, and duration), participants had faster (better) performance on a mobile Stroop-like task of executive function (mCWIT). This relationship persisted after adjusting for HIV status and other demographics (age and race or ethnicity). No effects of PA were found on a task of verbal learning (mVLT). The results are consistent with much of the laboratory-based literature linking an acute increase in PA to improvements in executive functions [11-13] and extending it to provide novel evidence that these associations may be captured remotely in the real world using EMCTs and accelerometry.

Next, we examined whether the significant cross-sectional association of daily PA with EMCT performance varied as a function of cognitive status, functional status (IADL), and cardiovascular risk, all of which elevate the risk for ADRD [43-47]. We found that only for individuals who reported being functionally dependent on IADL, there was a significant association between daily PA and executive function performance. It has been suggested that individuals who are at higher risk of ADRD may benefit most from lifestyle interventions to reduce cognitive decline [25]. Our findings are in line with this literature, suggesting that those with more functional limitations may see a greater cognitive benefit from engagement in PA in their free-living environments. We did

XSL•FO
RenderX

not find moderation effects related to HIV, cognitive, or cardiovascular risk status, suggesting that the daily relationship of PA with executive function does not significantly vary as a result of these risk factors in this limited sample of adults. Importantly, the association of daily PA with daily executive function performance persisted after adjusting these models for HIV status, global cognition, and cardiovascular risk, reinforcing the robustness of this relationship in a heterogeneous sample.

The findings suggest that daily variation in PA is sensitive to fluctuations in daily executive function performance, which can be reliably measured remotely using actigraphy and EMCTs. This has important implications for the development of novel digital health interventions to lower the risk of ADRD. Given the many limitations preventing individuals from participating in supervised or group-based PAs (ie, transportation barriers, safety issues, and mobility limitations) and the high costs associated with in-person interventions, it is important to develop scalable, low-cost, evidence-based [48] interventions to promote PA in real-world contexts [49]. Digital health technologies can help us achieve this goal by tracking intervention adherence and providing feedback in real time [50]. Future lifestyle interventions to preserve brain health can leverage these technologies to measure PA and cognition in real-world contexts to better understand treatment effects, improve generalizability, and reach out to a larger sector of the population who may not otherwise be able to attend laboratory-based (in-person) clinical trial visits. Moreover, digital health technologies can help track intervention adherence and changes in outcomes of interest, such as cognition, which can be measured repeatedly over the course of a clinical trial rather than only at baseline and post intervention time points to better capture change over time.

Limitations

A limitation of this study is its observational rather than interventional nature, which does not allow us to determine whether improvements in PA lead to better executive function performance or vice versa. Second, the measurement time was limited to 14 days, which may have limited the range of variability in PA and cognitive testing performance. Third, this study was performed on a small and heterogeneous sample of adults with various comorbid conditions, limiting our ability to adjust the models for other potential confounders that may have affected EMCT performance, such as mood and motivation indicators. That said, it is unlikely that our results would change if the models were adjusted for mood, given previous findings showing that adjusting for mood did not alter the significant association of daily activities with EMCT performance in this sample [51]. Moreover, given the small sample size, null findings may have resulted from a lack of power rather than an absence of associations. Fourth, accelerometers were worn on the wrist rather than the hip, which has been traditionally thought to provide a more accurate measurement of PA [52], although this assumption has recently been challenged, given the ubiquitous nature of wrist-worn accelerometry and their improved estimation of PA [40,53,54]. Fifth, participants were paid to complete the EMCTs; therefore, adherence may differ in studies where participants do not receive financial compensation for completing the assessments. Finally, future studies with larger sample sizes should account for more variables that may affect the relationship between PA and cognition, such as the length and severity of HIV disease.

Conclusions

In conclusion, using EMCTs and accelerometry to capture cognitive performance and PA in free-living environments may be an ecologically valid means of capturing real-world associations between cognition and PA. Our findings suggest that these digital techniques are promising candidates for tracking cognitive change and may be useful in the context of lifestyle (nonpharmacological) digital interventions designed to reduce ADRD risk and improve brain and cognitive health. Having participants complete cognitive tests in more familiar settings, such as their home or during other daily activities, can help increase the generalizability of findings, reduce intervention costs, increase scalability, and improve adherence to digital health interventions.

Acknowledgments

This work was supported by the National Institutes of Health (RCM: grant numbers National Institute of Mental Health [NIMH] K23MH105297, NIMH K23 MH107260 S1, and National Institute on Aging R01AG062387), LMC F31 AG067869, and ZZZ: grant numbers NIA K23AG049906 and R01AG066657. The study was more broadly supported by the HIV Neurobehavioral Research Center (HNRC) award P30MH062512 from NIMH.

The HNRC is supported by the center award P30MH062512 from NIMH. The San Diego HNRC group is affiliated with the University of California, San Diego; the Naval Hospital, San Diego; and the Veterans Affairs San Diego Healthcare System, and includes Robert K Heaton, PhD (director); Igor Grant, MD (codirector); Ronald J Ellis, MD, PhD, and Scott Letendre, MD (associate directors); Jennifer Iudicello, PhD (center manager); Donald Franklin, Jr; Melanie Sherman and Cheryl Kelley; NeuroMedical core members, including Ronald J Ellis, MD, PhD (director and neuromedical unit head), Scott Letendre, MD (coinvestigator and laboratory unit head), and Christine Fennema-Notestine, PhD (coinvestigator and neuroimaging unit head); Debra Rosario, MPH; Neurobehavioral and Psychiatry core members, including David J Moore, PhD (codirector and neurobehavioral unit head), Murray B Stein, MD (codirector and psychiatry unit head), Erin E Morgan, PhD (coinvestigator and psychiatric coordinator), Andrew H Miller, and Matthew Dawson; NeuroVirology and Biology core members, including Sara Gianella Weibel, MD (codirector and neurovirology unit head), Sarah A LaMere, DVM, PhD (associate unit head, neurovirology unit), Cristian Achim, MD, PhD (codirector and neurobiology unit head), Ana Sanchez, PhD (coinvestigator and neurobiology unit), and Adam Fields, PhD (associate unit head, neurobiology unit); Microbiome core members, including Rob Knight, PhD

XSL•FO
RenderX

(codirector) and Pieter Dorrestein, PhD (codirector); Developmental core members, including Scott Letendre, MD (director), Ajay Bharti, MD (coinvestigator), J Allen McCutchan, MD, and Christine Fennema-Notestine, PhD; Administrative core members, including Robert K Heaton, PhD (director and coordinating unit head); Participant Accrual and Retention Unit members, including J Hampton Atkinson, MD (unit head) and Jennifer Marquie-Beck, MPH; Data Management and Information Systems Unit members, including Ian Abramson, PhD (unit head) and Clint Cushman; and Statistics Unit members, including Florin Vaida, PhD (unit head), Anya Umlauf, MS, and BT, PhD. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the US government. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

RCM is a cofounder of KeyWise, Inc and a consultant for NeuroUX. The terms of these arrangements have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies.

Multimedia Appendix 1
HIV Neurobehavioral Research Program neuropsychology battery.
[DOCX File , 20 KB - mhealth_v10i1e33747_app1.docx]

References

1. 2020 Alzheimer's disease facts and figures. Alzheimers Dement 2020 Mar;16(3):391-460 (forthcoming). [doi: 10.1002/alz.12068] [Medline: 32157811]
2. Dhana K, Evans DA, Rajan KB, Bennett DA, Morris MC. Healthy lifestyle and the risk of Alzheimer dementia: findings from 2 longitudinal studies. Neurology 2020 Jul 28;95(4):e374-e383 [FREE Full text] [doi: 10.1212/WNL.0000000000009816] [Medline: 32554763]
3. Combination of healthy lifestyle traits may substantially reduce Alzheimer's. National Institutes of Health. URL: https://www.nih.gov/news-events/news-releases/combination-healthy-lifestyle-traits-may-substantially-reduce-alzheimers [accessed 2022-01-05]
4. Paganini-Hill A, Kawas C, Corrada M. Lifestyle factors and dementia in the oldest-old: the 90+ study. Alzheimer Dis Assoc Disord 2016;30(1):21-26 [FREE Full text] [doi: 10.1097/WAD.0000000000000087] [Medline: 25710250]
5. Blondell SJ, Hammersley-Mather R, Veerman JL. Does physical activity prevent cognitive decline and dementia?: a systematic review and meta-analysis of longitudinal studies. BMC Public Health 2014 May 27;14:510 [FREE Full text] [doi: 10.1186/1471-2458-14-510] [Medline: 24885250]
6. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. J Intern Med 2011 Jan;269(1):107-117 [FREE Full text] [doi: 10.1111/j.1365-2796.2010.02281.x] [Medline: 20831630]
7. Sanders LM, Hortobágyi T, la Bastide-van Gemert S, van der Zee EA, van Heuvelen MJ. Dose-response relationship between exercise and cognitive function in older adults with and without cognitive impairment: a systematic review and meta-analysis. PLoS One 2019;14(1):e0210036 [FREE Full text] [doi: 10.1371/journal.pone.0210036] [Medline: 30629631]
8. Spartano NL, Demissie S, Himali JJ, Dukes KA, Murabito JM, Vasan RS, et al. Accelerometer-determined physical activity and cognitive function in middle-aged and older adults from two generations of the Framingham Heart Study. Alzheimers Dement (N Y) 2019;5:618-626 [FREE Full text] [doi: 10.1016/j.trci.2019.08.007] [Medline: 31660424]
9. Kennedy G, Hardman RJ, Macpherson H, Scholey AB, Pipingas A. How does exercise reduce the rate of age-associated cognitive decline? A review of potential mechanisms. J Alzheimers Dis 2017;55(1):1-18. [doi: 10.3233/JAD-160665] [Medline: 27636853]
10. Ratey J, Loehr J. The positive impact of physical activity on cognition during adulthood: a review of underlying mechanisms, evidence and recommendations. Rev Neurosci 2011;22(2):171-185. [doi: 10.1515/RNS.2011.017] [Medline: 21417955]
11. Chen F, Etnier J, Wu C, Cho Y, Hung T, Chang Y. Dose-response relationship between exercise duration and executive function in older adults. J Clin Med 2018 Sep 13;7(9):279 [FREE Full text] [doi: 10.3390/jcm7090279] [Medline: 30217031]
12. Kamijo K, Hayashi Y, Sakai T, Yahiro T, Tanaka K, Nishihira Y. Acute effects of aerobic exercise on cognitive function in older adults. J Gerontol B Psychol Sci Soc Sci 2009 May;64(3):356-363. [doi: 10.1093/geronb/gbp030] [Medline: 19363089]
13. Lefferts WK, DeBlois JP, White CN, Heffernan KS. Effects of acute aerobic exercise on cognition and constructs of decision-making in adults with and without hypertension. Front Aging Neurosci 2019;11:41 [FREE Full text] [doi: 10.3389/fnagi.2019.00041] [Medline: 30906257]
14. Voss M, Weng T, Narayana-Kumanan K, Cole RC, Wharff C, Reist L, et al. Acute exercise effects predict training change in cognition and connectivity. Med Sci Sports Exerc 2020 Jan;52(1):131-140 [FREE Full text] [doi: 10.1249/MSS.0000000000002115] [Medline: 31385912]

XSL•FO
RenderX

15. Loprinzi P, Blough J, Crawford L, Ryu S, Zou L, Li H. The temporal effects of acute exercise on episodic memory function: systematic review with meta-analysis. Brain Sci 2019 Apr 18;9(4):87 [FREE Full text] [doi: 10.3390/brainsci9040087] [Medline: 31003491]
16. Chang Y, Labban J, Gapin J, Etnier J. The effects of acute exercise on cognitive performance: a meta-analysis. Brain Res 2012 May 09;1453:87-101. [doi: 10.1016/j.brainres.2012.02.068] [Medline: 22480735]
17. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038 [FREE Full text] [doi: 10.1093/geroni/igy038] [Medline: 30619948]
18. Moore R, Campbell L, Delgadillo J, Paolillo EW, Sundermann EE, Holden J, et al. Smartphone-based measurement of executive function in older adults with and without HIV. Arch Clin Neuropsychol 2020 May 22;35(4):347-357 [FREE Full text] [doi: 10.1093/arclin/acz084] [Medline: 31942632]
19. Moore RC, Paolillo EW, Sundermann EE, Campbell LM, Delgadillo J, Heaton A, et al. Validation of the mobile verbal learning test: illustration of its use for age and disease-related cognitive deficits. Int J Methods Psychiatr Res 2021 Mar;30(1):e1859 [FREE Full text] [doi: 10.1002/mpr.1859] [Medline: 33159488]
20. Parrish E, Kamarsu S, Harvey P, Pinkham A, Depp C, Moore R. Remote ecological momentary testing of learning and memory in adults with serious mental illness. Schizophr Bull 2021 Apr 29;47(3):740-750 [FREE Full text] [doi: 10.1093/schbul/sbaa172] [Medline: 33219382]
21. Weizenbaum E, Torous J, Fulford D. Cognition in context: understanding the everyday predictors of cognitive performance in a new era of measurement. JMIR Mhealth Uhealth 2020 Jul 23;8(7):e14328 [FREE Full text] [doi: 10.2196/14328] [Medline: 32706680]
22. Yuan J, Libon DJ, Karjadi C, Ang AF, Devine S, Auerbach SH, et al. Association between the digital clock drawing test and neuropsychological test performance: large community-based prospective cohort (Framingham heart study). J Med Internet Res 2021 Jun 08;23(6):e27407 [FREE Full text] [doi: 10.2196/27407] [Medline: 34100766]
23. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2022-01-05]
24. Older adults keep pace on tech usage. AARP. URL: https://www.aarp.org/research/topics/technology/info-2019/2020-technology-trends-older-americans.html [accessed 2022-01-05]
25. Hussenoeder FS, Riedel-Heller SG. Primary prevention of dementia: from modifiable risk factors to a public brain health agenda? Soc Psychiatry Psychiatr Epidemiol 2018 Dec;53(12):1289-1301. [doi: 10.1007/s00127-018-1598-7] [Medline: 30255384]
26. Wilkinson G, Robertson G. Wide Range Achievement Test. United Kingdom: Pearson; 2006.
27. Paolillo EW, Obermeit LC, Tang B, Depp CA, Vaida F, Moore DJ, et al. Smartphone-based ecological momentary assessment (EMA) of alcohol and cannabis use in older adults with and without HIV infection. Addict Behav 2018 Aug;83:102-108 [FREE Full text] [doi: 10.1016/j.addbeh.2017.10.016] [Medline: 29126667]
28. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]
29. Heaton RK, Clifford DB, Franklin DR, Woods SP, Ake C, Vaida F, CHARTER Group. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. Neurology 2010 Dec 07;75(23):2087-2096 [FREE Full text] [doi: 10.1212/WNL.0b013e318200d727] [Medline: 21135382]
30. Cysique LA, Franklin D, Abramson I, Ellis RJ, Letendre S, Collier A, CHARTER Group, HNRC Group. Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. J Clin Exp Neuropsychol 2011 Jun;33(5):505-522 [FREE Full text] [doi: 10.1080/13803395.2010.535504] [Medline: 21391011]
31. Heaton R, Taylor M, Manly J. Chapter 5 - Demographic Effects and Use of Demographically Corrected Norms with the WAIS-III and WMS-III. In: Clinical Interpretation of the WAIS-III and WMS-III. Cambridge, Massachusetts, United States: Academic Press; 2003:181-210.
32. Heaton RK, Marcotte TD, Mindt MR, Sadek J, Moore DJ, Bentley H, HNRC Group. The impact of HIV-associated neuropsychological impairment on everyday functioning. J Int Neuropsychol Soc 2004 May;10(3):317-331. [doi: 10.1017/S1355617704102130] [Medline: 15147590]
33. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
34. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743-753. [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]
35. Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L, Experts Panel. Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil 2010 Apr;17(2):127-139. [doi: 10.1097/HJR.0b013e32832ed875] [Medline: 20215971]
36. Robusto KM, Trost SG. Comparison of three generations of ActiGraph™ activity monitors in children and adolescents. J Sports Sci 2012;30(13):1429-1435 [FREE Full text] [doi: 10.1080/02640414.2012.710761] [Medline: 22857599]

XSL•FO
RenderX

37. John D, Freedson P. ActiGraph and Actical physical activity monitors: a peek under the hood. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S86-S89 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399f5e] [Medline: 22157779]
38. Matthews CE, Hagströmer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S68-S76 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399e5b] [Medline: 22157777]
39. Freedson P, Bowles H, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S1-S4 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399b7e] [Medline: 22157769]
40. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]
41. Tudor-Locke C, Barreira TV, Schuna JM, Mire EF, Chaput J, Fogelholm M, ISCOLE Research Group. Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). Int J Behav Nutr Phys Act 2015 Feb 11;12:11 [FREE Full text] [doi: 10.1186/s12966-015-0172-x] [Medline: 25881074]
42. Bassett D, Troiano R, McClain J, Wolff D. Accelerometer-based physical activity: total volume per day and standardized measures. Med Sci Sports Exerc 2015 Apr;47(4):833-838. [doi: 10.1249/MSS.0000000000000468] [Medline: 25102292]
43. Agüero-Torres H, Fratiglioni L, Guo Z, Viitanen M, von Strauss E, Winblad B. Dementia is the major cause of functional dependence in the elderly: 3-year follow-up data from a population-based study. Am J Public Health 1998 Oct;88(10):1452-1456. [doi: 10.2105/ajph.88.10.1452] [Medline: 9772843]
44. Fillit H, Nash DT, Rundek T, Zuckerman A. Cardiovascular risk factors and dementia. Am J Geriatr Pharmacother 2008 Jun;6(2):100-118. [doi: 10.1016/j.amjopharm.2008.06.004] [Medline: 18675769]
45. Michaud T, Su D, Siahpush M, Murman D. The risk of incident mild cognitive impairment and progression to dementia considering mild cognitive impairment subtypes. Dement Geriatr Cogn Dis Extra 2017;7(1):15-29 [FREE Full text] [doi: 10.1159/000452486] [Medline: 28413413]
46. Sabia S, Fayosse A, Dumurgier J, Schnitzler A, Empana J, Ebmeier KP, et al. Association of ideal cardiovascular health at age 50 with incidence of dementia: 25 year follow-up of Whitehall II cohort study. BMJ 2019 Aug 07;366:l4414 [FREE Full text] [doi: 10.1136/bmj.l4414] [Medline: 31391187]
47. Visser PJ, Kester A, Jolles J, Verhey F. Ten-year risk of dementia in subjects with mild cognitive impairment. Neurology 2006 Oct 10;67(7):1201-1207. [doi: 10.1212/01.wnl.0000238517.59286.c5] [Medline: 17030753]
48. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Effects of three motivationally targeted mobile device applications on initial physical activity and sedentary behavior change in midlife and older adults: a randomized trial. PLoS One 2016;11(6):e0156370 [FREE Full text] [doi: 10.1371/journal.pone.0156370] [Medline: 27352250]
49. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. Exp Gerontol 2019 Jun;120:68-87. [doi: 10.1016/j.exger.2019.02.020] [Medline: 30836130]
50. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
51. Campbell LM, Paolillo EW, Heaton A, Tang B, Depp CA, Granholm E, et al. Daily activities related to mobile cognitive performance in middle-aged and older adults: an ecological momentary cognitive assessment study. JMIR Mhealth Uhealth 2020 Sep 24;8(9):e19579 [FREE Full text] [doi: 10.2196/19579] [Medline: 32969829]
52. Chen KY, Acra SA, Majchrzak K, Donahue CL, Baker L, Clemens L, et al. Predicting energy expenditure of physical activity using hip- and wrist-worn accelerometers. Diabetes Technol Ther 2003;5(6):1023-1033 [FREE Full text] [doi: 10.1089/152091503322641088] [Medline: 14709206]
53. Kamada M, Shiroma EJ, Harris TB, Lee I. Comparison of physical activity assessed using hip- and wrist-worn accelerometers. Gait Posture 2016 Feb;44:23-28 [FREE Full text] [doi: 10.1016/j.gaitpost.2015.11.005] [Medline: 27004628]
54. Kerr J, Marinac CR, Ellis K, Godbole S, Hipp A, Glanz K, et al. Comparison of accelerometry methods for estimating physical activity. Med Sci Sports Exerc 2017 Mar;49(3):617-624 [FREE Full text] [doi: 10.1249/MSS.0000000000001124] [Medline: 27755355]

Abbreviations

ADRD: Alzheimer's disease and related dementias
CPM: counts per minute
CVD: cardiovascular disease
EMCT: ecological momentary cognitive test
HNRC: HIV Neurobehavioral Research Center
HNRP: HIV Neurobehavioral Research Program
IADL: instrumental activities of daily living

XSL•FO
RenderX

mCWIT: mobile color–word interference test
mVLT: mobile verbal learning test
NIMH: National Institute of Mental Health
PA: physical activity
PWH: people living with HIV

Edited by L Buis; submitted 21.09.21; peer-reviewed by L Smith, E Weizenbaum, P Harvey; comments to author 10.11.21; revised version received 23.11.21; accepted 09.12.21; published 31.01.22.

Please cite as:
Zlatar ZZ, Campbell LM, Tang B, Gabin S, Heaton A, Higgins M, Swendsen J, Moore DJ, Moore RC
Daily Level Association of Physical Activity and Performance on Ecological Momentary Cognitive Tests in Free-living Environments: A Mobile Health Observational Study
JMIR Mhealth Uhealth 2022;10(1):e33747
URL: https://mhealth.jmir.org/2022/1/e33747
doi:10.2196/33747
PMID:

©Zvinka Z Zlatar, Laura M Campbell, Bin Tang, Spenser Gabin, Anne Heaton, Michael Higgins, Joel Swendsen, David J Moore, Raeanne C Moore. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 31.01.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX