

---

# JMIR mHealth and uHealth

---

Impact Factor (2022): 5.0  
Volume 9 (2021), Issue 4 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

---

## Contents

### Viewpoint

- The Potential of mHealth as a Game Changer for the Management of Sickle Cell Disease in India ([e25496](#))  
Ravindra Kumar, Aparup Das. . . . . 5

### Reviews

- Diagnostic Accuracy of Ambulatory Devices in Detecting Atrial Fibrillation: Systematic Review and Meta-analysis ([e26167](#))  
Tien Yang, Li Huang, Shwetambara Malwade, Chien-Yi Hsu, Yang Chen. . . . . 12
- Biofeedback-Based Connected Mental Health Interventions for Anxiety: Systematic Literature Review ([e26038](#))  
Mahra Alneyadi, Nidal Drissi, Mariam Almeqbaali, Sofia Ouhbi. . . . . 27
- Impact of Mobile Health Devices for the Detection of Atrial Fibrillation: Systematic Review ([e26161](#))  
Tom Biersteker, Martin Schalij, Roderick Treskes. . . . . 40
- Barriers to Use of Remote Monitoring Technologies Used to Support Patients With COVID-19: Rapid Review ([e24743](#))  
Elizabeth Houlding, Kedar Mate, Kim Engler, David Ortiz-Paredes, Marie-Pascale Pomey, Joseph Cox, Tarek Hijal, Bertrand Lebouché. . . . . 416

### Original Papers

- Accuracy of Smart Scales on Weight and Body Composition: Observational Study ([e22487](#))  
Justine Frija-Masson, Jimmy Mullaert, Emmanuelle Vidal-Petiot, Nathalie Pons-Kerjean, Martin Flamant, Marie-Pia d'Ortho. . . . . 53
- Noninvasive Hemoglobin Level Prediction in a Mobile Phone Environment: State of the Art Review and Recommendations ([e16806](#))  
Md Hasan, Md Aziz, Md Zarif, Mahmudul Hasan, MMA Hashem, Shion Guha, Richard Love, Sheikh Ahamed. . . . . 61
- Staff Perceptions of Preimplementation Barriers and Facilitators to a Mobile Health Antiretroviral Therapy Adherence Counseling Intervention in South Africa: Qualitative Study ([e23280](#))  
Siobhan McCreesh-Toselli, John Torline, Hetta Gouse, Reuben Robbins, Claude Mellins, Robert Remien, Jessica Rowe, Neshaan Peton, Stephan Rabie, John Joska. . . . . 85

**Influences on the Uptake of Health and Well-being Apps and Curated App Portals: Think-Aloud and Interview Study (e27173)**  
 Dorothy Szinay, Olga Perski, Andy Jones, Tim Chadborn, Jamie Brown, Felix Naughton. . . . . 97

**Leprosy Screening Based on Artificial Intelligence: Development of a Cross-Platform App (e23718)**  
 Márcio De Souza, Gabriel Lopes, Alexandre Branco, Jessica Fairley, Lucia Fraga. . . . . 113

**Shaping Workflows in Digital and Remote Diabetes Care During the COVID-19 Pandemic via Service Design: Prospective, Longitudinal, Open-label Feasibility Trial (e24374)**  
 Katarina Braune, Karina Boss, Jessica Schmidt-Herzel, Katarzyna Gajewska, Axel Thieffry, Lilian Schulze, Barbara Posern, Klemens Raile. . . . . 1  
 2 6

**A Patient-Oriented App (ThessHF) to Improve Self-Care Quality in Heart Failure: From Evidence-Based Design to Pilot Study (e24271)**  
 Constantinos Bakogiannis, Anastasios Tsarouchas, Dimitrios Mouselimis, Charalampos Lazaridis, Efstratios Theofillogianakos, Antonios Billis, Stergios Tzikas, Nikolaos Fragakis, Panagiotis Bamidis, Christodoulos Papadopoulos, Vassilios Vassilikos. . . . . 145

**Usability of a Mobile App for Real-Time Assessment of Fatigue and Related Symptoms in Patients With Multiple Sclerosis: Observational Study (e19564)**  
 Miklos Palotai, Max Wallack, Gergo Kujbus, Adam Dalnoki, Charles Guttmann. . . . . 159

**Slip Buddy App for Weight Management: Randomized Feasibility Trial of a Dietary Lapse Tracking App (e24249)**  
 Sherry Pagoto, Bengisu Tulu, Molly Waring, Jared Goetz, Jessica Bibeau, Joseph Divito, Laurie Groshon, Matthew Schroeder. . . . . 173

**Quality, Features, and Presence of Behavior Change Techniques in Mobile Apps Designed to Improve Physical Activity in Pregnant Women: Systematic Search and Content Analysis (e23649)**  
 Melanie Hayman, Kristie-Lee Alfrey, Summer Cannon, Stephanie Alley, Amanda Rebar, Susan Williams, Camille Short, Abby Altazan, Natalie Comardelle, Sinead Currie, Caitlin Denton, Cheryce Harrison, Tayla Lamerton, Gabriela Mena, Lisa Moran, Michelle Mottola, Taniya Nagpal, Lisa Vincze, Stephanie Schoeppe. . . . . 190

**Usability and Acceptability of a Mobile App for the Self-Management of Alcohol Misuse Among Veterans (Step Away): Pilot Cohort Study (e25927)**  
 Carol Malte, Patrick Dulin, John Baer, John Fortney, Anissa Danner, Aline Lott, Eric Hawkins. . . . . 202

**Mobile Phone Intervention Based on an HIV Risk Prediction Tool for HIV Prevention Among Men Who Have Sex With Men in China: Randomized Controlled Trial (e19511)**  
 Ke Yun, Zhenxing Chu, Jing Zhang, Wenqing Geng, Yongjun Jiang, Willa Dong, Hong Shang, Junjie Xu. . . . . 218

**Feasibility and Preliminary Efficacy of a Community-Based Addiction Rehabilitation Electronic System in Substance Use Disorder: Pilot Randomized Controlled Trial (e21087)**  
 Xiaomin Xu, Shujuan Chen, Junning Chen, Zhikang Chen, Liming Fu, Dingchen Song, Min Zhao, Haifeng Jiang. . . . . 229

**Development of the Niggle App for Supporting Young People on Their Dynamic Journey to Well-being: Co-design and Qualitative Research Study (e21085)**  
 Stoyan Stoyanov, Oksana Zelenko, Aleksandra Staneva, David Kavanagh, Calvin Smith, Gavin Sade, Jessica Cheers, Leanne Hides. . . . . 244

**Results of a Culturally Tailored Smartphone-Delivered Physical Activity Intervention Among Midlife African American Women: Feasibility Trial (e27383)**  
 Rodney Joseph, Barbara Ainsworth, Kevin Hollingshead, Michael Todd, Colleen Keller. . . . . 261

**Evaluation of Mood Check-in Feature for Participation in Meditation Mobile App Users: Retrospective Longitudinal Analysis (e27106)**  
 Jennifer Huberty, Jeni Green, Megan Puzia, Chad Stecher. . . . . 279

<b>An Implementation Strategy to Expand Mobile Health Use in HIV Care Settings: Rapid Evaluation Study Using the Consolidated Framework for Implementation Research (e19163)</b>	
Wendy Cohn, Chelsea Canan, Sarah Knight, Ava Waldman, Rebecca Dillingham, Karen Ingersoll, Julie Schexnayder, Tabor Flickinger. . . . .	288
<b>Using a Mobile Phone App to Analyze the Relationship Between Planned and Performed Physical Activity in University Students: Observational Study (e17581)</b>	
Matthew Stewart, Taylor Nezich, Joyce Lee, Rebecca Hasson, Natalie Colabianchi. . . . .	303
<b>Self-Efficacy and Motivation to Quit of Smokers Seeking to Quit: Quantitative Assessment of Smoking Cessation Mobile Apps (e25030)</b>	
Nikita Rajani, Nikolaos Mastellos, Filippos Filippidis. . . . .	316
<b>Text Messaging Intervention for Young Smokers Experiencing Homelessness: Lessons Learned From a Randomized Controlled Trial (e23989)</b>	
Sebastian Linnemayr, Rushil Zutshi, William Shadel, Eric Pedersen, Maria DeYoreo, Joan Tucker. . . . .	328
<b>Gender Differences in Satisfaction With a Text Messaging Program (Text4Hope) and Anticipated Receptivity to Technology-Based Health Support During the COVID-19 Pandemic: Cross-sectional Survey Study (e24184)</b>	
Reham Shalaby, Wesley Vuong, Marianne Hrabok, April Gusnowski, Kelly Mrklas, Daniel Li, Mark Snaterse, Shireen Surood, Bo Cao, Xin-Min Li, Russell Greiner, Andrew Greenshaw, Vincent Agyapong. . . . .	338
<b>Designing an App to Overcome Language Barriers in the Delivery of Emergency Medical Services: Participatory Development Process (e21586)</b>	
Eva Noack, Jennifer Schulze, Frank Müller. . . . .	354
<b>Baby Buddy App for Breastfeeding and Behavior Change: Retrospective Study of the App Using the Behavior Change Wheel (e25668)</b>	
Loretta Musgrave, Alison Baum, Nilushka Perera, Caroline Homer, Adrienne Gordon. . . . .	368
<b>Smartphone Users' Persuasion Knowledge in the Context of Consumer mHealth Apps: Qualitative Study (e16518)</b>	
Eunsin Joo, Anastasia Kononova, Shaheen Kanthawala, Wei Peng, Shelia Cotten. . . . .	378
<b>Acceptance, Barriers, and Future Preferences of Mobile Health Among Patients Receiving Trauma and Orthopedic Surgical Care: Paper-Based Survey in a Prospective Multicenter Study (e23784)</b>	
Felix Reinecke, Florian Dittrich, Marcel Dudda, Andreas Stang, Christina Polan, Roman Müller, Paula Beck, Max Kauther. . . . .	393
<b>Voice Interface Technology Adoption by Patients With Heart Failure: Pilot Comparison Study (e24646)</b>	
Lida Apergi, Margret Bjarnadottir, John Baras, Bruce Golden, Kelley Anderson, Jiling Chou, Nawar Shara. . . . .	401
<b>A Smartphone App Designed to Empower Patients to Contribute Toward Safer Surgical Care: Qualitative Evaluation of Diverse Public and Patient Perceptions Using Focus Groups (e24065)</b>	
Stephanie Russ, Nick Sevdalis, Josephine Ocloo. . . . .	432
<b>Measuring the Quality of Clinical Skills Mobile Apps for Student Learning: Systematic Search, Analysis, and Comparison of Two Measurement Scales (e25377)</b>	
Tehmina Gladman, Grace Tylee, Steve Gallagher, Jonathan Mair, Rebecca Grainger. . . . .	447
<b>Effects of App-Based Transitional Care on the Self-Efficacy and Quality of Life of Patients With Spinal Cord Injury in China: Randomized Controlled Trial (e22960)</b>	
Ting Liu, Sumei Xie, Yingmin Wang, Jie Tang, Xiaokuo He, Tiebin Yan, Kun Li. . . . .	459
<b>Remote Heart Rhythm Monitoring by Photoplethysmography-Based Smartphone Technology After Cardiac Surgery: Prospective Observational Study (e26519)</b>	
Marie Lamberigts, Lucas Van Hoof, Tine Proesmans, Pieter Vandervoort, Lars Grieten, Peter Haemers, Filip Rega. . . . .	480

Perception of Purposeful and Recreational Smartphone Use in Physiotherapy: Randomized Controlled Trial ( <a href="#">e25717</a> ) Martina Bientzle, Anne Restle, Joachim Kimmerle. ....	490
Assessing the Quality of Mobile Health-Related Apps: Interrater Reliability Study of Two Guides ( <a href="#">e26471</a> ) Jordi Miró, Pere Llorens-Vernet. ....	503
Twelve-Lead Electrocardiogram Acquisition With a Patchy-Type Wireless Device in Ambulance Transport: Simulation-Based Randomized Controlled Trial ( <a href="#">e24142</a> ) Sunyoung Yoon, Taerim Kim, Taehwan Roh, Hansol Chang, Sung Hwang, Hee Yoon, Tae Shin, Min Sim, Ik Jo, Won Cha. ....	512
Force-Sensitive Mat for Vertical Jump Measurement to Assess Lower Limb Strength: Validity and Reliability Study ( <a href="#">e27336</a> ) Erik Vanegas, Yolocauhtli Salazar, Raúl Igual, Inmaculada Plaza. ....	522
Relationship Between Major Depression Symptom Severity and Sleep Collected Using a Wristband Wearable Device: Multicenter Longitudinal Observational Study ( <a href="#">e24604</a> ) Yuezhou Zhang, Amos Folarin, Shaoxiong Sun, Nicholas Cummins, Rebecca Bendayan, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Petroula Laiou, Faith Matcham, Katie White, Femke Lamers, Sara Siddi, Sara Simblett, Inez Myin-Germeys, Aki Rintala, Til Wykes, Josep Haro, Brenda Penninx, Vaibhav Narayan, Matthew Hotopf, Richard Dobson, RADAR-CNS Consortium. ....	537
Associations of Physical Activity Level and Variability With 6-Month Weight Change Among 26,935 Users of Connected Devices: Observational Real-Life Study ( <a href="#">e25385</a> ) Douae El Fatouhi, Lidia Delrieu, Catherine Goetzinger, Laurent Malisoux, Aurélie Affret, David Campo, Guy Fagherazzi. ....	552

## Corrigenda and Addenda

Correction: A Novel Remote Follow-Up Tool Based on an Instant Messaging/Social Media App for the Management of Patients With Low Anterior Resection Syndrome: Pilot Prospective Self-Control Study ( <a href="#">e29325</a> ) Fan Liu, Peng Guo, Xiangqian Su, Ming Cui, Jianlong Jiang, Suo Wang, Zhouman Yu, Runhe Zhou, Yingjiang Ye. ....	567
--	-----

Viewpoint

# The Potential of mHealth as a Game Changer for the Management of Sickle Cell Disease in India

Ravindra Kumar<sup>1</sup>, PhD; Aparup Das<sup>1</sup>, PhD

ICMR-National Institute of Research in Tribal Health, Jabalpur, India

**Corresponding Author:**

Aparup Das, PhD

ICMR-National Institute of Research in Tribal Health

ICMR-NIRTH Campus

Nagpur Road, PO Garha

Jabalpur, 482003

India

Phone: 91 7612370800

Email: [aparupdas@nirth.res.in](mailto:aparupdas@nirth.res.in)

## Abstract

Sickle cell disease (SCD) is a chronic genetic disease that requires lifelong therapy and monitoring. Low drug adherence and poor monitoring may lead to an increase in morbidities and low quality of life. In the era of digital technology, various mobile health (mHealth) apps are being tested for their potential in increasing drug adherence in patients with SCD. We herewith discuss the applicability and feasibility of these mHealth apps for the management of SCD in India.

(*JMIR Mhealth Uhealth* 2021;9(4):e25496) doi:[10.2196/25496](https://doi.org/10.2196/25496)

**KEYWORDS**

sickle cell disease; drug adherence; mHealth; India

## Background

In recent years, a revolution in information technologies has greatly influenced health care practices under the broad definition of digital health. Digital health practices are becoming highly adaptable in both developed and developing countries [1]. Moreover, with the increasing use of smartphones, smart watches, and artificial intelligence-based devices, mobile health (mHealth) is expected to define the standard of health care delivery across the globe. Specifically, mHealth apps can be used to increase disease awareness, increase drug adherence, provide cognitive behavioral therapies, and track health care delivery [2-5]. There are more than 325,000 mHealth apps available for Android and Apple smartphones [6]. Various mHealth apps have been clinically tested for their effect on compliance for many chronic diseases worldwide [7-11] and there are mounting indications that support the feasibility and applicability of mHealth interventions for better compliance in managing chronic diseases in pediatric patients as well. During the ongoing COVID-19 pandemic, mHealth has emerged as a silver bullet, not only for teleconsultations and telemonitoring of patients with chronic diseases, but also for increasing health care delivery in remote areas [12-14].

## Sickle Cell Disease: A Life-threatening and Highly Morbid Disorder

Sickle cell disease (SCD) is a genetic and chronic ailment, highly prevalent in Sub-Saharan Africa, the Middle East, the Mediterranean region, India, and parts of Central and South America. Globally, more than 300,000 children are born each year with SCD and three countries (Nigeria, India, and the Democratic Republic of Congo) bear about half of the global burden [15]. Patients with SCD often present with acute complications (eg, bone pain crisis, acute abdominal pain, acute chest syndrome, visceral sequestration crisis, aplastic crisis, acute anemia, cerebrocardiovascular complications, priapism). Chronic morbidities in SCD (eg, chronic pain syndromes, immunological and infectious complications, chronic lung disease, hepatobiliary complications, renal complications, leg ulcer, musculoskeletal complications, and psychosocial or psychiatric issues) are often encountered [16].

The under-five mortality of SCD varies significantly depending on the availability of health care facilities and infrastructure. For example, in low-income countries with poor access to health care services, mortality can reach up to 90% [15]. There is now growing evidence that continuous interventions through disease-modifying drugs such as hydroxyurea and prophylactic

antibiotics can decrease morbidities and increase life expectancy, thereby leading to increased health-related quality of life and reduced health burden [17-22]. However, the sustainability of drug adherence is a major challenge in public health, as an average of about 50% of patients with chronic illness (including SCD) do not adhere to proper treatment in developed countries [23,24]. Furthermore, the prevalence of medication nonadherence is much higher in developing countries due to their relative scarcity and inequities of health care resources in comparison to developed nations [25-27]. Low drug adherence not only is associated with increased morbidity and poor health-related quality of life for patients, but also increases burden on the health economy and increases health care utilization in a country setting.

Various factors govern drug adherence, including delivery of health care services, economic situation, and cultural factors of patients [28]. Behavioral factors like forgetfulness, inappropriate time management, lack of awareness of disease, and fear of drugs pose additional barriers to drug adherence [29-31]. It has been observed that parents of patients with SCD with mild symptoms are less willing to accept the risk associated with taking hydroxyurea, particularly with regard to the long-term side effects, which include birth defects and cancer [29]. Therefore, patients with mild symptoms are more unwilling to take medications. In addition, patients with poor drug adherence can experience a mild-to-moderate medication response, causing frustration in both patients and their parents, which further leads to increased drug nonadherence.

Low drug adherence in SCD is also a major issue in India, where the majority of patients with SCD belong to scheduled communities (both the scheduled caste and scheduled tribes) [32,33], who face many barriers in accessing quality health care services. The scheduled caste and scheduled tribes are groups of people scheduled in the Constitution of India on the basis of economic, social, and educational disadvantage. Tribes are the most marginalized communities and mostly live in hard-to-reach remote hilly and forested areas. Many of the tribal communities are still dependent upon hunting and gathering and primitive agricultural practices. Poor accessibility, social disconnection, inconvenient timing, longer waiting times in government health care facilities, and poor economic conditions are some of the major roadblocks to accessing quality health care services in tribal areas in India [34]. Further, due to the different environments and terrains in which tribes live, inequalities in sociocultural behavior, and lack of participation, a universal design of health care services becomes inappropriate in tribal communities. Various nongovernmental organizations (NGOs)—such as SEARCH (in Gadchiroli, Maharashtra, India) [35], MAHAN (Melghat, Maharashtra) [36], Seva Rural (Jhagadia, Gujarat) [37], Jan Swasthya Sahyog (Ganiyari, Chhattisgarh) [38], and ASHWINI (Gudalur, Tamil Nadu) [39]—have developed different locally appropriate customized models for providing better health care services in hard-to-reach remote tribal areas.

With regard to SCD, various control programs have been in operation in India for an extended period of time. However, due to the lack of an organizational referral system and standard treatment guidelines for the management of SCD, most of these

programs are limited to the screening of patients [40]. It has been observed that most patients seek treatment only when they have an acute sickle cell crisis. Moreover, the low accessibility and affordability of drugs is a major obstacle in the management of SCD in Indian tribes [41,42].

## *The Role of mHealth in SCD Management: Indian Context*

Disease interventions using mHealth have been tried for several diseases in tribal and rural India. For example, in a recent clinical trial, accredited social health activists (ASHAs) who used an mHealth app (ImTecho) as a job aid were able to provide better maternal and child health services in tribal areas in the state of Gujarat [43]. Similarly, presumptive tuberculosis referrals increased when rural health care providers used mHealth technology in tribal areas of Khuntu District of Jharkhand state [44]. Based on different models, a framework of digital health for increasing referrals, monitoring patients, and increasing the accessibility of malaria drugs in rural areas has been suggested [45]. A similar approach can be adopted for increasing drug availability and improving the referral system for SCD as well. Intervention measures (eg, sending reminders, allowing pain and symptom reporting, enabling self-management, and providing cognitive behavioral therapy) through mHealth apps for increasing drug adherence for SCD have been tested in several parts of the globe and have shown promising results in terms of disease outcome [46-48]. In addition, self-management practices in SCD have been found to be increased by the use of mHealth apps. However, no such intervention measures have previously been tried for patients with SCD in India. Considering SCD in India is mostly prevalent in scheduled communities living in rural and hard-to-reach areas, mHealth might be especially useful for SCD management. This is because access to health care services in hard-to-reach rural and forested areas remains meagre due to poor local transport systems. In these circumstances, mHealth apps can be used to improve the referral system in inaccessible areas by increasing knowledge among ASHAs, community health workers, and the traditional healers in the tribal communities.

Is it practically feasible to use mHealth apps for patients with SCD in India? Several broad factors might limit the adoption of mHealth in patients with SCD in India (Textbox 1). One of the major roadblocks is poor internet availability in rural India. Although the number of mobile and internet users in rural areas is said to have increased substantially in recent years (Figure 1) [49], by the year 2018, just 14.9% of rural households had internet access [50]. Furthermore, mobile and internet connectivity are considered poor in hard-to-reach forested areas, which could hinder access to mHealth apps. In addition, the poor socioeconomic status of tribes further limits their ability to sustain the cost of smartphones and internet data plans for a longer duration. Moreover, a low level of digital literacy among rural and tribal people also impedes the usability of mHealth apps. Apart from this, people living in rural areas, especially tribes, have their own social beliefs and customs that are different from other populations in India, thereby limiting the use of such mHealth apps due to hesitation and stigma. In

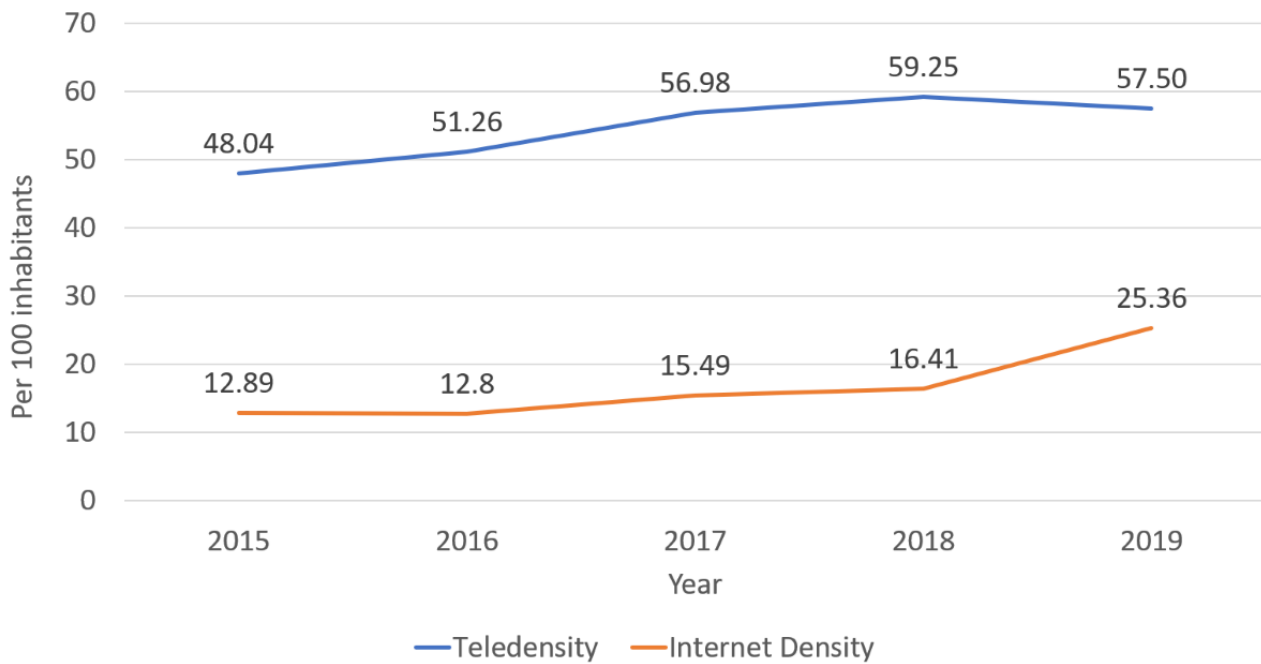
addition, privacy and cybersecurity issues related to the online use of any mobile app (including mHealth) may be a concern of people who are inadequately digitally literate and economically disadvantaged. Further, most mHealth apps require

manual use; therefore, after a certain time period, patient engagement may decrease. Therefore, not only the feasibility, but also the long-term sustainability of mHealth apps in patients with SCD residing in rural and tribal areas is presently uncertain.

**Textbox 1.** Utility of and roadblocks to mobile health apps for the management of sickle cell disease in India.

Utility
<ul style="list-style-type: none"> <li>• Enhancement of referral system</li> <li>• Strengthens public health delivery</li> <li>• Augments drug adherence</li> <li>• Creates disease awareness</li> <li>• Improves self-management abilities during a primary sickle cell crisis</li> </ul>
Roadblocks
<ul style="list-style-type: none"> <li>• Low digital literacy</li> <li>• Poor telecommunication connectivity in remote areas</li> <li>• Cost of data plans for economically disadvantaged communities</li> <li>• Digital privacy and data security</li> </ul>

**Figure 1.** Telephone and internet density in rural India.



In spite of the above mentioned barriers to the use of mHealth apps in India, there is a light at the end of the tunnel. Based on the applicability of disease management and growing digitalization in India, measures should be taken to implement mHealth technology for SCD interventions with locally appropriate customized solutions. For example, mHealth apps with artificial intelligence that require little manual intervention may be helpful for patients with low digital literacy. In addition, for better management of SCD, user-specific customizable mHealth apps should be made, not only for increasing drug adherence but also for informing the patient about disease precipitating factors through daily activity monitoring and environmental conditions, thereby reducing acute crises. Such

customization requires field-based and need-based approaches through direct inputs from patients as well as their caregivers in terms of expectations, needs, and experiences in combatting SCD to improve short-term engagement. Once such customized and user-friendly mHealth apps are in use by patients, regular feedback in terms of usage, technicality, and user-friendliness can be obtained from users at regular intervals for long-term sustainability. Furthermore, considering the poor economic conditions of many rural and tribal people, the provision of incentives (in term of free smartphones and internet data packs) may also serve as a boon for increased sustainability of mHealth apps for SCD intervention in India. In addition, encouragement through community engagement as part of a drive for increased

use of mHealth apps could prove useful, as could the involvement of health ambassadors (eg, ASHAs, community health workers, community leaders, traditional healers, and/or adult SCD champions). These health ambassadors can be trained on the operation and use of mHealth apps, who in turn can train patients with SCD and their caregivers to use mHealth apps effectively. Moreover, socioeducational drives for increasing digital literacy among patients with SCD residing in rural and tribal areas can prove highly rewarding in terms of effective management. Furthermore, studies aiming to evaluate the cost effectiveness of mHealth interventions in improving health care delivery and drug adherence for patients with SCD living in remote areas are needed to assess the sustainability of such interventions.

### *The Way Forward*

In 2016, the Indian government aimed to increase the accessibility of hydroxyurea in district hospitals located in high-prevalence areas and issued guidelines for the prevention and control of hemoglobinopathies [51]. However, due to implementation gaps, the program could not be initiated in the majority of states, particularly in tribal areas [40]. Recently, a draft policy was prepared, which advocates for the improvement of SCD treatment centers in all districts and ensures a free supply of hydroxyurea and penicillin to low-income patients [52]. Furthermore, the recent introduction of the Ayushman Bharat scheme promulgating universal health coverage ensures free treatment of low-income patients admitted in hospitals [53]. For the successful implementation of these schemes, there is a pressing need for the development of an organizational referral

system and a strong communication system for increasing awareness about diseases and the treatment and monitoring of patients. In this context, mHealth can appropriately establish itself as a game changer in the management of SCD in India. Therefore, the introduction of mHealth for SCD has huge potential in terms of enhancing health care delivery and management, and providing a better quality of life to patients globally, including in India. However, the feasibility, acceptability, and sustainability of such mHealth apps for the management of SCD in India is uncertain at present. For this, the keys to success are the following: a strong inclination from the government through health policy reform for socially and economically marginalized populations living in rural and hard-to-reach areas; a significant investment in infrastructure development for the strengthening of mobile and internet connectivity in these areas; and promulgation of digital literacy using intersectoral coordination and public-private partnerships in rural and tribal populations. The recent introduction of the Prime Minister's Digital India Movement and the Prime Minister's Rural Digital Literacy Movement—with a view to ensuring the availability of cost-effective high-speed internet to every citizen and empowering the rural population in the use of digital technologies, including the marginalized scheduled castes/tribes and differently abled persons—is one such welcome move in this direction. Furthermore, with advances in digital technologies, a reduction in the cost of internet data plans and a higher penetration of mobile and internet connectivity in rural and hard-to-reach areas are expected in the future. In light of the above developments, the future of mHealth in India seems bright, and this will lead to a better prognosis and improved health-related quality of life for patients with SCD.

### **Acknowledgments**

The authors thank the Secretary, Department of Health Research and the Director General, Indian Council of Medical Research (ICMR) for providing encouragement and facilities. The authors are also thankful to Dr Nishant Saxena and Dr Anil Kumar Verma, ICMR-NIRTH, Jabalpur, for providing valuable suggestions. Critical comments from three anonymous reviewers helped us improve the manuscript. The manuscript has been approved by the Publication Screening Committee of ICMR-NIRTH, Jabalpur, and assigned the number ICMR-NIRTH/PSC/06/2021.

### **Conflicts of Interest**

None declared.

### **References**

1. Radbron E, Wilson V, McCance T, Middleton R. The Use of Data Collected From mHealth Apps to Inform Evidence-Based Quality Improvement: An Integrative Review. *Worldviews Evid Based Nurs* 2019 Feb;16(1):70-77. [doi: [10.1111/wvn.12343](https://doi.org/10.1111/wvn.12343)] [Medline: [30758133](https://pubmed.ncbi.nlm.nih.gov/30758133/)]
2. Hossain MM, Tasnim S, Sharma R, Sultana A, Shaik AF, Faizah F, et al. Digital interventions for people living with non-communicable diseases in India: A systematic review of intervention studies and recommendations for future research and development. *Digit Health* 2019;5:2055207619896153 [FREE Full text] [doi: [10.1177/2055207619896153](https://doi.org/10.1177/2055207619896153)] [Medline: [31897307](https://pubmed.ncbi.nlm.nih.gov/31897307/)]
3. Xu H, Long H. The Effect of Smartphone App-Based Interventions for Patients With Hypertension: Systematic Review and Meta-Analysis. *JMIR mHealth uHealth* 2020 Oct 19;8(10):e21759 [FREE Full text] [doi: [10.2196/21759](https://doi.org/10.2196/21759)] [Medline: [33074161](https://pubmed.ncbi.nlm.nih.gov/33074161/)]
4. 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](https://doi.org/10.2196/13608)] [Medline: [31512580](https://pubmed.ncbi.nlm.nih.gov/31512580/)]



5. Psihogios AM, Stiles-Shields C, Neary M. The Needle in the Haystack: Identifying Credible Mobile Health Apps for Pediatric Populations during a Pandemic and beyond. *J Pediatr Psychol* 2020 Nov 01;45(10):1106-1113 [[FREE Full text](#)] [doi: [10.1093/jpepsy/jsaa094](https://doi.org/10.1093/jpepsy/jsaa094)] [Medline: [33068424](https://pubmed.ncbi.nlm.nih.gov/33068424/)]
6. Larson RS. A Path to Better-Quality mHealth Apps. *JMIR mHealth uHealth* 2018 Jul 30;6(7):e10414 [[FREE Full text](#)] [doi: [10.2196/10414](https://doi.org/10.2196/10414)] [Medline: [30061091](https://pubmed.ncbi.nlm.nih.gov/30061091/)]
7. Yang Y, Chen H, Qazi H, Morita PP. Intervention and Evaluation of Mobile Health Technologies in Management of Patients Undergoing Chronic Dialysis: Scoping Review. *JMIR mHealth uHealth* 2020 Apr;8(4):e15549 [[FREE Full text](#)] [doi: [10.2196/15549](https://doi.org/10.2196/15549)] [Medline: [32242823](https://pubmed.ncbi.nlm.nih.gov/32242823/)]
8. Kompala T, Neinstein AB. Telehealth in type 1 diabetes. *Curr Opin Endocrinol Diabetes Obes* 2021 Feb 01;28(1):21-29. [doi: [10.1097/MED.0000000000000600](https://doi.org/10.1097/MED.0000000000000600)] [Medline: [33332927](https://pubmed.ncbi.nlm.nih.gov/33332927/)]
9. Agarwal RK, Sedai A, Ankita K, Parmar L, Dhanya R, Dhimal S, et al. Information Technology-Assisted Treatment Planning and Performance Assessment for Severe Thalassemia Care in Low- and Middle-Income Countries: Observational Study. *JMIR Med Inform* 2019 Jan 23;7(1):e9291 [[FREE Full text](#)] [doi: [10.2196/medinform.9291](https://doi.org/10.2196/medinform.9291)] [Medline: [30672740](https://pubmed.ncbi.nlm.nih.gov/30672740/)]
10. Choi H, Van Riper M. mHealth Family Adaptation Intervention for Families of Young Children with Down Syndrome: A Feasibility Study. *J Pediatr Nurs* 2020;50:e69-e76. [doi: [10.1016/j.pedn.2019.03.010](https://doi.org/10.1016/j.pedn.2019.03.010)] [Medline: [30928311](https://pubmed.ncbi.nlm.nih.gov/30928311/)]
11. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. *Psychooncology* 2020 Jan;29(1):17-37. [doi: [10.1002/pon.5280](https://doi.org/10.1002/pon.5280)] [Medline: [31692183](https://pubmed.ncbi.nlm.nih.gov/31692183/)]
12. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of Telehealth During the COVID-19 Pandemic: Scoping Review. *J Med Internet Res* 2020 Dec 01;22(12):e24087. [doi: [10.2196/24087](https://doi.org/10.2196/24087)] [Medline: [33147166](https://pubmed.ncbi.nlm.nih.gov/33147166/)]
13. Andrews E, Berghofer K, Long J, Prescott A, Caboral-Stevens M. Satisfaction with the use of telehealth during COVID-19: An integrative review. *Int J Nurs Stud Adv* 2020 Nov;2:100008 [[FREE Full text](#)] [doi: [10.1016/j.ijnsa.2020.100008](https://doi.org/10.1016/j.ijnsa.2020.100008)] [Medline: [33083791](https://pubmed.ncbi.nlm.nih.gov/33083791/)]
14. Badawy SM, Radovic A. Digital Approaches to Remote Pediatric Health Care Delivery During the COVID-19 Pandemic: Existing Evidence and a Call for Further Research. *JMIR Pediatr Parent* 2020 Jun 25;3(1):e20049 [[FREE Full text](#)] [doi: [10.2196/20049](https://doi.org/10.2196/20049)] [Medline: [32540841](https://pubmed.ncbi.nlm.nih.gov/32540841/)]
15. Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anaemia in children under five, 2010-2050: modelling based on demographics, excess mortality, and interventions. *PLoS Med* 2013 Jul;10(7):e1001484 [[FREE Full text](#)] [doi: [10.1371/journal.pmed.1001484](https://doi.org/10.1371/journal.pmed.1001484)] [Medline: [23874164](https://pubmed.ncbi.nlm.nih.gov/23874164/)]
16. Piel FB, Steinberg MH, Rees DC. Sickle Cell Disease. *N Engl J Med* 2017 Apr 20;376(16):1561-1573. [doi: [10.1056/NEJMra1510865](https://doi.org/10.1056/NEJMra1510865)] [Medline: [28423290](https://pubmed.ncbi.nlm.nih.gov/28423290/)]
17. Candrilli SD, O'Brien SH, Ware RE, Nahata MC, Seiber EE, Balkrishnan R. Hydroxyurea adherence and associated outcomes among Medicaid enrollees with sickle cell disease. *Am J Hematol* 2011 Mar;86(3):273-277 [[FREE Full text](#)] [doi: [10.1002/ajh.21968](https://doi.org/10.1002/ajh.21968)] [Medline: [21328441](https://pubmed.ncbi.nlm.nih.gov/21328441/)]
18. Badawy SM, Thompson AA, Lai J, Penedo FJ, Rychlik K, Liem RI. Adherence to hydroxyurea, health-related quality of life domains, and patients' perceptions of sickle cell disease and hydroxyurea: a cross-sectional study in adolescents and young adults. *Health Qual Life Outcomes* 2017 Jul 05;15(1):136 [[FREE Full text](#)] [doi: [10.1186/s12955-017-0713-x](https://doi.org/10.1186/s12955-017-0713-x)] [Medline: [28679417](https://pubmed.ncbi.nlm.nih.gov/28679417/)]
19. Badawy SM, Thompson AA, Holl JL, Penedo FJ, Liem RI. Healthcare utilization and hydroxyurea adherence in youth with sickle cell disease. *Pediatr Hematol Oncol* 2018;35(5-6):297-308. [doi: [10.1080/08880018.2018.1505988](https://doi.org/10.1080/08880018.2018.1505988)] [Medline: [30636474](https://pubmed.ncbi.nlm.nih.gov/30636474/)]
20. Walsh KE, Cutrona SL, Kavanagh PL, Crosby LE, Malone C, Lobner K, et al. Medication adherence among pediatric patients with sickle cell disease: a systematic review. *Pediatrics* 2014 Dec;134(6):1175-1183 [[FREE Full text](#)] [doi: [10.1542/peds.2014-0177](https://doi.org/10.1542/peds.2014-0177)] [Medline: [25404717](https://pubmed.ncbi.nlm.nih.gov/25404717/)]
21. Loiselle K, Lee JL, Szulcowski L, Drake S, Crosby LE, Pai ALH. Systematic and Meta-Analytic Review: Medication Adherence Among Pediatric Patients With Sickle Cell Disease. *J Pediatr Psychol* 2016 May;41(4):406-418 [[FREE Full text](#)] [doi: [10.1093/jpepsy/jsv084](https://doi.org/10.1093/jpepsy/jsv084)] [Medline: [26384715](https://pubmed.ncbi.nlm.nih.gov/26384715/)]
22. Shih S, Cohen LL. A Systematic Review of Medication Adherence Interventions in Pediatric Sickle Cell Disease. *J Pediatr Psychol* 2020 Jul 01;45(6):593-606. [doi: [10.1093/jpepsy/jsaa031](https://doi.org/10.1093/jpepsy/jsaa031)] [Medline: [32417887](https://pubmed.ncbi.nlm.nih.gov/32417887/)]
23. Sabaté E. Adherence to long-term therapies: evidence for action. Geneva, Switzerland: World Health Organization; 2003. URL: [https://www.who.int/chp/knowledge/publications/adherence\\_full\\_report.pdf?ua=1](https://www.who.int/chp/knowledge/publications/adherence_full_report.pdf?ua=1) [accessed 2020-12-28]
24. Fernandez-Lazaro CI, García-González JM, Adams DP, Fernandez-Lazaro D, Mielgo-Ayuso J, Caballero-García A, et al. Adherence to treatment and related factors among patients with chronic conditions in primary care: a cross-sectional study. *BMC Fam Pract* 2019 Sep 14;20(1):132 [[FREE Full text](#)] [doi: [10.1186/s12875-019-1019-3](https://doi.org/10.1186/s12875-019-1019-3)] [Medline: [31521114](https://pubmed.ncbi.nlm.nih.gov/31521114/)]
25. Santra G. Assessment of adherence to cardiovascular medicines in rural population: An observational study in patients attending a tertiary care hospital. *Indian J Pharmacol* 2015;47(6):600-604 [[FREE Full text](#)] [doi: [10.4103/0253-7613.169573](https://doi.org/10.4103/0253-7613.169573)] [Medline: [26729949](https://pubmed.ncbi.nlm.nih.gov/26729949/)]

26. Saqlain M, Riaz A, Malik MN, Khan S, Ahmed A, Kamran S, et al. Medication Adherence and Its Association with Health Literacy and Performance in Activities of Daily Livings among Elderly Hypertensive Patients in Islamabad, Pakistan. *Medicina (Kaunas)* 2019 May 18;55(5):163 [FREE Full text] [doi: [10.3390/medicina55050163](https://doi.org/10.3390/medicina55050163)] [Medline: [31109105](https://pubmed.ncbi.nlm.nih.gov/31109105/)]
27. Nonogaki A, Heang H, Yi S, van Pelt M, Yamashina H, Taniguchi C, et al. Factors associated with medication adherence among people with diabetes mellitus in poor urban areas of Cambodia: A cross-sectional study. *PLoS One* 2019 Nov 19;14(11):e0225000 [FREE Full text] [doi: [10.1371/journal.pone.0225000](https://doi.org/10.1371/journal.pone.0225000)] [Medline: [31743349](https://pubmed.ncbi.nlm.nih.gov/31743349/)]
28. McQuaid EL, Landier W. Cultural Issues in Medication Adherence: Disparities and Directions. *J Gen Intern Med* 2018 Feb;33(2):200-206 [FREE Full text] [doi: [10.1007/s11606-017-4199-3](https://doi.org/10.1007/s11606-017-4199-3)] [Medline: [29204971](https://pubmed.ncbi.nlm.nih.gov/29204971/)]
29. Badawy SM, Thompson AA, Liem RI. Beliefs about hydroxyurea in youth with sickle cell disease. *Hematol Oncol Stem Cell Ther* 2018 Sep;11(3):142-148 [FREE Full text] [doi: [10.1016/j.hemonc.2018.01.001](https://doi.org/10.1016/j.hemonc.2018.01.001)] [Medline: [29397333](https://pubmed.ncbi.nlm.nih.gov/29397333/)]
30. Badawy SM, Shah R, Beg U, Heneghan MB. Habit Strength, Medication Adherence, and Habit-Based Mobile Health Interventions Across Chronic Medical Conditions: Systematic Review. *J Med Internet Res* 2020 Apr 28;22(4):e17883. [doi: [10.2196/17883](https://doi.org/10.2196/17883)]
31. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. *Pediatr Blood Cancer* 2016 May;63(5):848-852. [doi: [10.1002/pbc.25905](https://doi.org/10.1002/pbc.25905)] [Medline: [26844685](https://pubmed.ncbi.nlm.nih.gov/26844685/)]
32. Hockham C, Bhatt S, Colah R, Mukherjee MB, Penman BS, Gupta S, et al. The spatial epidemiology of sickle-cell anaemia in India. *Sci Rep* 2018 Dec 06;8(1):17685 [FREE Full text] [doi: [10.1038/s41598-018-36077-w](https://doi.org/10.1038/s41598-018-36077-w)] [Medline: [30523337](https://pubmed.ncbi.nlm.nih.gov/30523337/)]
33. Colah R, Mukherjee M, Ghosh K. Sickle cell disease in India. *Curr Opin Hematol* 2014 May;21(3):215-223. [doi: [10.1097/MOH.000000000000029](https://doi.org/10.1097/MOH.000000000000029)] [Medline: [24714525](https://pubmed.ncbi.nlm.nih.gov/24714525/)]
34. Narain JP. Health of tribal populations in India: How long can we afford to neglect? *Indian J Med Res* 2019 Mar;149(3):313-316 [FREE Full text] [doi: [10.4103/ijmr.IJMR\\_2079\\_18](https://doi.org/10.4103/ijmr.IJMR_2079_18)] [Medline: [31249192](https://pubmed.ncbi.nlm.nih.gov/31249192/)]
35. Bang AT, Reddy HM, Deshmukh MD, Baitule SB, Bang RA. Neonatal and infant mortality in the ten years (1993 to 2003) of the Gadchiroli field trial: effect of home-based neonatal care. *J Perinatol* 2005 Mar;25 Suppl 1:S92-107. [doi: [10.1038/sj.jp.7211277](https://doi.org/10.1038/sj.jp.7211277)] [Medline: [15791283](https://pubmed.ncbi.nlm.nih.gov/15791283/)]
36. Report of MAHAN Trust 1997-2019. URL: <https://www.mahantrust.org/policy-changes> [accessed 2020-12-28]
37. Shah SP, Shah P, Desai S, Modi D, Desai G, Arora H. Effectiveness and Feasibility of Weekly Iron and Folic Acid Supplementation to Adolescent Girls and Boys through Peer Educators at Community Level in the Tribal Area of Gujarat. *Indian J Community Med* 2016;41(2):158-161 [FREE Full text] [doi: [10.4103/0970-0218.173498](https://doi.org/10.4103/0970-0218.173498)] [Medline: [27051093](https://pubmed.ncbi.nlm.nih.gov/27051093/)]
38. Jain Y, Kataria R, Patil S, Kadam S, Kataria A, Jain R, et al. *Indian J Med Res* 2015 May;141(5):663-672 [FREE Full text] [doi: [10.4103/0971-5916.159582](https://doi.org/10.4103/0971-5916.159582)] [Medline: [26139787](https://pubmed.ncbi.nlm.nih.gov/26139787/)]
39. Nimgaonkar V, Krishnamurti L, Prabhakar H, Menon N. Comprehensive integrated care for patients with sickle cell disease in a remote aboriginal tribal population in southern India. *Pediatr Blood Cancer* 2014 Apr;61(4):702-705. [doi: [10.1002/pbc.24723](https://doi.org/10.1002/pbc.24723)] [Medline: [24347362](https://pubmed.ncbi.nlm.nih.gov/24347362/)]
40. Geethakumari K, Kusuma YS, Babu BV. Beyond the screening: The need for health systems intervention for prevention and management of sickle cell disease among tribal population of India. *Int J Health Plann Manage* 2021 Mar;36(2):236-243. [doi: [10.1002/hpm.3081](https://doi.org/10.1002/hpm.3081)] [Medline: [33000499](https://pubmed.ncbi.nlm.nih.gov/33000499/)]
41. Jain D, Lothe A, Colah R. Sickle Cell Disease: Current Challenges. *J Hematol Thrombo Dis* 2015;03(06):224. [doi: [10.4172/2329-8790.1000224](https://doi.org/10.4172/2329-8790.1000224)]
42. Odame I, Jain D. *Indian J Med Res* 2020 Jun;151(6):505-508 [FREE Full text] [doi: [10.4103/ijmr.IJMR\\_2064\\_20](https://doi.org/10.4103/ijmr.IJMR_2064_20)] [Medline: [32719221](https://pubmed.ncbi.nlm.nih.gov/32719221/)]
43. Modi D, Dholakia N, Gopalan R, Venkatraman S, Dave K, Shah S, et al. mHealth intervention "ImTeCHO" to improve delivery of maternal, neonatal, and child care services-A cluster-randomized trial in tribal areas of Gujarat, India. *PLoS Med* 2019 Oct;16(10):e1002939 [FREE Full text] [doi: [10.1371/journal.pmed.1002939](https://doi.org/10.1371/journal.pmed.1002939)] [Medline: [31647821](https://pubmed.ncbi.nlm.nih.gov/31647821/)]
44. Chadha S, Trivedi A, Nagaraja SB, Sagili K. Using mHealth to enhance TB referrals in a tribal district of India. *Public Health Action* 2017 Jun 21;7(2):123-126 [FREE Full text] [doi: [10.5588/pha.16.0080](https://doi.org/10.5588/pha.16.0080)] [Medline: [28695085](https://pubmed.ncbi.nlm.nih.gov/28695085/)]
45. Nema S, Verma AK, Tiwari A, Bharti PK. Digital Health Care Services to Control and Eliminate Malaria in India. *Trends Parasitol* 2021 Feb;37(2):96-99. [doi: [10.1016/j.pt.2020.11.002](https://doi.org/10.1016/j.pt.2020.11.002)] [Medline: [33262008](https://pubmed.ncbi.nlm.nih.gov/33262008/)]
46. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. *Pediatr Blood Cancer* 2016 May;63(5):848-852. [doi: [10.1002/pbc.25905](https://doi.org/10.1002/pbc.25905)] [Medline: [26844685](https://pubmed.ncbi.nlm.nih.gov/26844685/)]
47. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-Centered eHealth Interventions for Children, Adolescents, and Adults With Sickle Cell Disease: Systematic Review. *J Med Internet Res* 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: [10.2196/10940](https://doi.org/10.2196/10940)] [Medline: [30026178](https://pubmed.ncbi.nlm.nih.gov/30026178/)]
48. Hankins JS, Shah N. Tackling adherence in sickle cell disease with mHealth. *Lancet Haematol* 2020 Oct;7(10):e713-e714. [doi: [10.1016/S2352-3026\(20\)30299-4](https://doi.org/10.1016/S2352-3026(20)30299-4)] [Medline: [32976750](https://pubmed.ncbi.nlm.nih.gov/32976750/)]
49. Economics Research Unit, Department of Telecommunications, Ministry of Communications, Government of India. *Telecom Statistics India 2019*. URL: <https://dot.gov.in/sites/default/files/Telecom%20Statistics%20India-2019.pdf?download=1> [accessed 2020-12-28]

50. Ministry of Statistics and Programme Implementation, Government of India. Household Social Consumption on Education in India. NSS 75th Round report. URL: [http://mospi.nic.in/sites/default/files/publication\\_reports/Report\\_585\\_75th\\_round\\_Education\\_final\\_1507\\_0.pdf](http://mospi.nic.in/sites/default/files/publication_reports/Report_585_75th_round_Education_final_1507_0.pdf) [accessed 2020-12-28]
51. National Health Mission, Ministry of Health and Family Welfare, Government of India. Prevention and Control of Hemoglobinopathies in India - Thalassemias, Sickle Disease and Other Variant Hemoglobins. Guidelines on Hemoglobinopathies in India. URL: [https://nhm.gov.in/images/pdf/programmes/RBSK/Resource\\_Documents/Guidelines\\_on\\_Hemoglobinopathies\\_in%20India.pdf](https://nhm.gov.in/images/pdf/programmes/RBSK/Resource_Documents/Guidelines_on_Hemoglobinopathies_in%20India.pdf) [accessed 2020-12-28]
52. Ministry of Health and Family Welfare, Government of India. Draft Policy for Prevention and Control of Hemoglobinopathies - Thalassemia, Sickle Cell Disease and Variant Hemoglobins in India. URL: <https://www.nhp.gov.in/NHPfiles/1.pdf> [accessed 2020-12-28]
53. Bhargava B, Paul V. Informing NCD control efforts in India on the eve of Ayushman Bharat. Lancet 2018 Sep 11:S0140-6736(18)32172-X [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32172-X](https://doi.org/10.1016/S0140-6736(18)32172-X)] [Medline: [30219331](https://pubmed.ncbi.nlm.nih.gov/30219331/)]

## Abbreviations

**ASHA:** accredited social health activist

**mHealth:** mobile health

**SCD:** sickle cell disease

*Edited by L Buis; submitted 04.11.20; peer-reviewed by A Bhatia, J Hankins, S Badawy; comments to author 21.11.20; revised version received 09.01.21; accepted 22.01.21; published 13.04.21.*

*Please cite as:*

*Kumar R, Das A*

*The Potential of mHealth as a Game Changer for the Management of Sickle Cell Disease in India*

*JMIR Mhealth Uhealth 2021;9(4):e25496*

URL: <https://mhealth.jmir.org/2021/4/e25496>

doi: [10.2196/25496](https://doi.org/10.2196/25496)

PMID: [33847598](https://pubmed.ncbi.nlm.nih.gov/33847598/)

©Ravindra Kumar, Aparup Das. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 13.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

# Diagnostic Accuracy of Ambulatory Devices in Detecting Atrial Fibrillation: Systematic Review and Meta-analysis

Tien Yun Yang<sup>1</sup>, MD; Li Huang<sup>2</sup>, MD; Shwetambara Malwade<sup>3</sup>, BHMS, MS; Chien-Yi Hsu<sup>4,5</sup>, MD; Yang Ching Chen<sup>2,6</sup>, MD, PhD

<sup>1</sup>School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

<sup>2</sup>Department of Family Medicine, Taipei Medical University Hospital, Taipei City, Taiwan

<sup>3</sup>International Center for Health Information Technology, Taipei Medical University, Taipei, Taiwan

<sup>4</sup>Division of Cardiology, Department of Internal Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

<sup>5</sup>Division of Cardiology and Cardiovascular Research Center, Department of Internal Medicine, Taipei Medical University Hospital, Taipei, Taiwan

<sup>6</sup>Department of Family Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei City, Taiwan

**Corresponding Author:**

Yang Ching Chen, MD, PhD

Department of Family Medicine

Taipei Medical University Hospital

No. 252, Wuxing St

Xinyi District

Taipei City, 110

Taiwan

Phone: 886 2 2737 2181 ext 3032

Email: [melisa26@tmu.edu.tw](mailto:melisa26@tmu.edu.tw)

## Abstract

**Background:** Atrial fibrillation (AF) is the most common cardiac arrhythmia worldwide. Early diagnosis of AF is crucial for preventing AF-related morbidity, mortality, and economic burden, yet the detection of the disease remains challenging. The 12-lead electrocardiogram (ECG) is the gold standard for the diagnosis of AF. Because of technological advances, ambulatory devices may serve as convenient screening tools for AF.

**Objective:** The objective of this review was to investigate the diagnostic accuracy of 2 relatively new technologies used in ambulatory devices, non-12-lead ECG and photoplethysmography (PPG), in detecting AF. We performed a meta-analysis to evaluate the diagnostic accuracy of non-12-lead ECG and PPG compared to the reference standard, 12-lead ECG. We also conducted a subgroup analysis to assess the impact of study design and participant recruitment on diagnostic accuracy.

**Methods:** This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. MEDLINE and EMBASE were systematically searched for articles published from January 1, 2015 to January 23, 2021. A bivariate model was used to pool estimates of sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and area under the summary receiver operating curve (SROC) as the main diagnostic measures. Study quality was evaluated using the quality assessment of diagnostic accuracy studies (QUADAS-2) tool.

**Results:** Our search resulted in 16 studies using either non-12-lead ECG or PPG for detecting AF, comprising 3217 participants and 7623 assessments. The pooled estimates of sensitivity, specificity, PLR, NLR, and diagnostic odds ratio for the detection of AF were 89.7% (95% CI 83.2%-93.9%), 95.7% (95% CI 92.0%-97.7%), 20.64 (95% CI 10.10-42.15), 0.11 (95% CI 0.06-0.19), and 224.75 (95% CI 70.10-720.56), respectively, for the automatic interpretation of non-12-lead ECG measurements and 94.7% (95% CI 93.3%-95.8%), 97.6% (95% CI 94.5%-99.0%), 35.51 (95% CI 18.19-69.31), 0.05 (95% CI 0.04-0.07), and 730.79 (95% CI 309.33-1726.49), respectively, for the automatic interpretation of PPG measurements.

**Conclusions:** Both non-12-lead ECG and PPG offered high diagnostic accuracies for AF. Detection employing automatic analysis techniques may serve as a useful preliminary screening tool before administering a gold standard test, which generally requires competent physician analyses. Subgroup analysis indicated variations of sensitivity and specificity between studies that recruited low-risk and high-risk populations, warranting future validity tests in the general population.

**Trial Registration:** PROSPERO International Prospective Register of Systematic Reviews CRD42020179937; [https://www.crd.york.ac.uk/prospero/display\\_record.php?RecordID=179937](https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=179937)

(*JMIR Mhealth Uhealth* 2021;9(4):e26167) doi:[10.2196/26167](https://doi.org/10.2196/26167)

## KEYWORDS

atrial fibrillation; ambulatory devices; electrocardiogram; photoplethysmography; diagnostic accuracy; ubiquitous health; mobile health; technology; ambulatory device

## Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia worldwide, affecting approximately 33.5 million individuals. AF is more prevalent with increasing age, and its prevalence is expected to double by 2030 [1]. The disease can result in considerable morbidity and mortality by increasing the risk of heart failure, stroke, major cardiovascular events, sudden cardiac death, chronic kidney disease, peripheral arterial disease, and all-cause mortality despite often being asymptomatic [2]. A range of management choices, including anticoagulation, rate control, and rhythm control through medication or electrical cardioversion, can markedly reduce risks and relieve symptoms.

Although the treatment of AF is well established because of numerous guidelines and clinical trials, the detection of the disease remains challenging. The gold standard for AF detection is the 12-lead electrocardiogram (ECG) [3]. However, this method is not always available and can be unfeasible in certain groups of patients, such as individuals with paroxysmal AF who fail to undergo a 12-lead ECG in a timely manner or individuals with silent (subclinical) AF without any related symptoms. Therefore, devices that are convenient, inexpensive, and ambulatory are required to serve as preliminary screening tools; subsequently, initial diagnoses can be confirmed or excluded using the gold standard 12-lead ECG in hospital settings [4,5].

Several studies have investigated the accuracy and applicability of various ambulatory devices over the past few years. Rather than focusing on the devices themselves, this review targeted technologies used in ambulatory devices; thus, the summarized results are not limited to certain products. Among reviews of the use of ambulatory devices for AF detection, only one systematic review and meta-analysis focused on the accuracy of technologies compared with the gold standard [6]. That review concluded that blood pressure monitors and non-12-lead ECGs were most accurate; however, it failed to consider a newer technology, photoplethysmography (PPG). PPG is a new technology that has become ubiquitous in recent years, and one of its most widely known implementations is the Apple Watch [7]. Therefore, we conducted an updated systematic review

focusing on 2 technologies that are used in ambulatory devices to detect AF: non-12-lead ECG and PPG. Additionally, the review focused on the automatic detection of AF utilizing built-in algorithms to validate their use as a convenient screening tool.

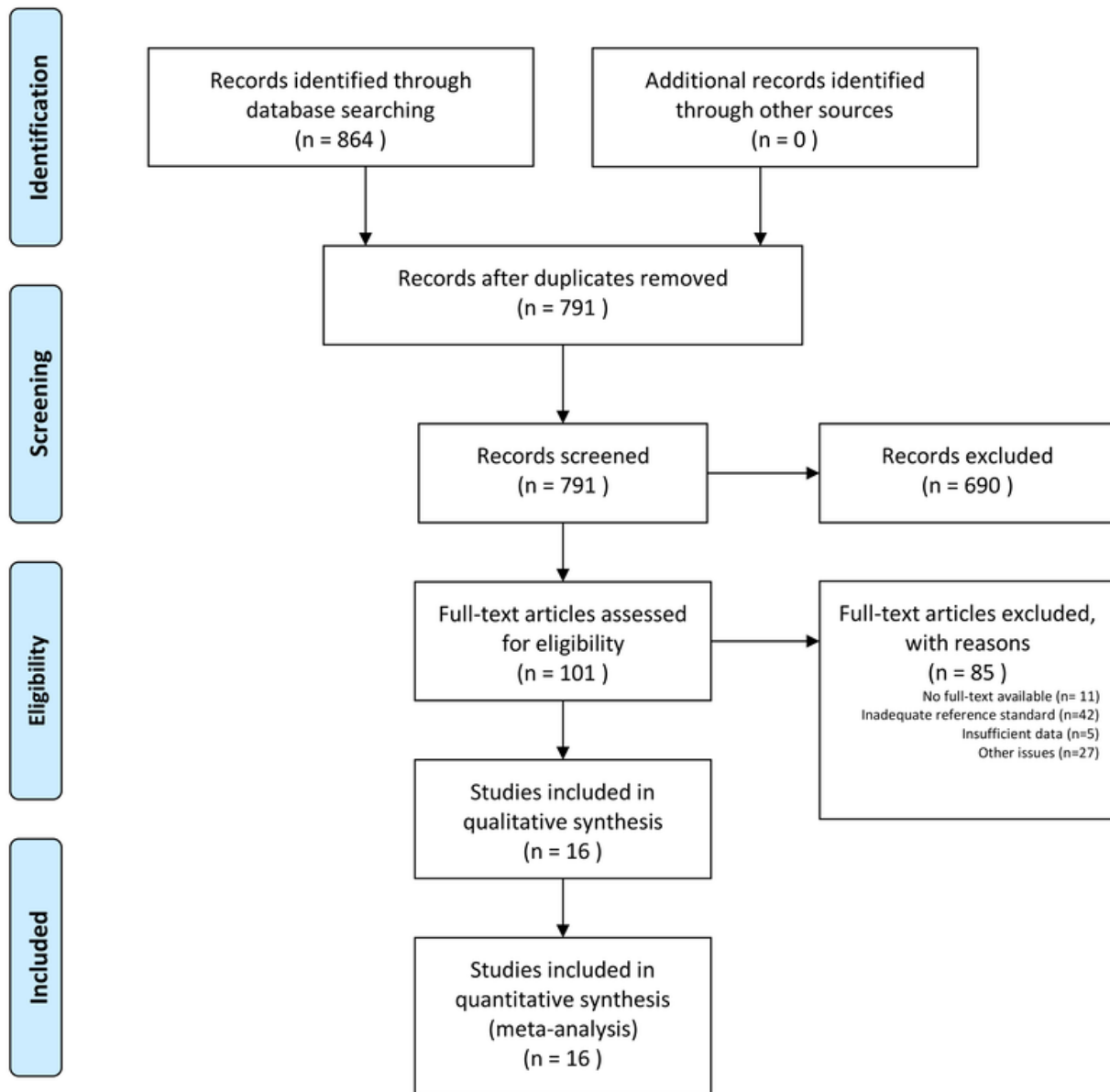
The aim of this paper was to provide a systematic overview of the accuracy of the 2 technologies compared with 12-lead ECG in the detection of AF as well as to describe their applicability, potential, and limitations.

## Methods

### Literature Search and Selection Criteria

We conducted a systematic search of MEDLINE (Ovid) and EMBASE for articles published from January 1, 2015 to January 23, 2021 using the search terms “mHealth,” “telemedicine,” “wearable,” “mobile health,” “mobile application,” and “digital treatment” in combination with the term “atrial fibrillation.” The search terms are presented in [Multimedia Appendix 1](#). In addition, reference lists of relevant systematic reviews and included studies were hand-searched to identify additional articles. Only papers in English were included. All randomized trials, observational studies, and case series were included, whereas systematic reviews and case reports were excluded. Studies that recruited participants aged  $\geq 18$  years, investigated any method of identifying patients with suspected AF using an ambulatory device equipped with automatic interpretation by a mobile app or algorithm, provided a reference standard with 12-lead ECG interpreted by a competent professional, and reported sufficient data to enable the calculation of the diagnostic accuracy were included. Studies that investigated invasive methods of identifying AF, focused on the training of algorithms, validated the method and algorithm through a dataset, or failed to provide a timely reference standard for all participants were excluded. Two reviewers independently conducted the screening and reviewing of articles, and any disagreements were resolved by consensus with a third reviewer. The study strictly followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, and the PRISMA search flow diagram is provided in [Figure 1](#).

**Figure 1.** Summary of the study selection process using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



**Data Extraction and Quality Assessment**

Two reviewers independently extracted the data for the selected articles, including the year of publication, country, study design, number of study participants, average age of the study population, characteristics of the study population, technology used for measurement, measurement time, and reference standard. The absolute numbers of true positive, false positive, false negative, and true negative were also extracted.

If a study provided data on both the measurement level and individual level, the data analysis performed on the measurement level was extracted first. If a study provided both training data and validation data, only validation data were extracted. If a study failed to provide sufficient data for the calculation of diagnostic accuracy, the lead authors of the studies were contacted to request the missing data. If the authors failed to

reply, 2 reviewers calculated the incomplete information on the basis of available data. Studies were excluded only if both the aforementioned methods failed to identify any additional data.

Study quality was evaluated using the quality assessment of diagnostic accuracy studies (QUADAS-2) tool, which is recommended for the evaluation of risk of bias and applicability in diagnostic accuracy studies [8,9]. The assessment tool focuses on 4 key domains: patient selection, index test, reference standard, and flow and timing. Risk of bias was evaluated in all 4 domains, and applicability concerns were evaluated in the first 3 domains. For each question, each study was graded as “low risk,” “high risk,” or “unclear risk.” A standardized table, recommended by the QUADAS-2 official website, was used to display the summarized results of the study quality appraisal.

## Statistical Analysis

Diagnostic accuracy statistics were computed using R software version 4.0.0 (R Core Team). The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) — with their respective 95% CIs — were calculated using a bivariate diagnostic random-effects model in the *mada* package [10]. Tests for heterogeneity regarding the DOR were also performed and presented with Cochran Q and Higgins I<sup>2</sup>. Summary receiver operating characteristic (SROC) curves were used to summarize and visualize the diagnostic performance of each included study. The area under the summary receiver operating characteristic (AUSROC) curve was also calculated for both the non-12-lead ECG and PPG methods.

Publication bias was assessed using the Deek funnel plot asymmetry test, which is a scatterplot of (1/root [ESS]) against (ln[DOR]) [11]; a *P* value <.10 indicated publication bias. Fagan nomogram analysis was performed to determine the posttest probability of the disease based on the likelihood ratio of the diagnostic test. An age-adjusted prevalence of 0.5% was applied in the analysis to represent the general population worldwide [12], whereas a prevalence of 2.3% was adopted to represent a population with higher risks as targeted in a systematic review [13]. The left axis of a Fagan nomogram represents the pretest probability, the middle axis displays the likelihood ratio of the diagnostic test, and the right axis indicates the posttest probability.

Because the risk of patient selection bias among the studies was obvious, a subgroup analysis between low-risk groups (studies that recruited patients with and without AF) and high-risk groups

(studies that recruited only patients with AF) was conducted to investigate the effect of the study design and population on diagnostic performance. Data synthesis and most statistical analyses were conducted using R version 4.0.0.

## Study Registration

This systematic review and meta-analysis was registered in the PROSPERO International Prospective Register of Systematic Reviews (ID: CRD42020179937).

## Results

### Literature Search

Figure 1 illustrates the flowchart of study inclusion. The initial search yielded 864 publications from MEDLINE (Ovid) and EMBASE. Duplicates and irrelevant studies were removed, yielding 791 publications for title and abstract review. After exclusion of 690 publications for irrelevant study focus, the rest of the 101 publications were then assessed through full-text articles. In total, 85 publications were excluded for the following reasons: having insufficient context, being an inappropriate study type, having a study population aged <18 years, utilizing invasive or implantable devices to identify AF, or having an inadequate reference standard. The full-text evaluation yielded 16 publications that met the inclusion criteria; these papers were included in this systematic review and meta-analysis [14-29]. No additional studies were identified through the hand-searching process. The characteristics of the included studies are listed in Table 1. The details of the study population, the prevalence of AF, ambulatory devices, measuring time, and measurement data are presented in 2 respective tables for non-12-lead ECG and PPG in Multimedia Appendix 2.

**Table 1.** Characteristics of the included studies and study population.

Study authors	Year	Country	Study design	Index test	n <sup>a</sup>	Population
Chen et al [14]	2020	China	Prospective, cross-sectional	Non-12-lead ECG <sup>b</sup> and PPG <sup>c</sup>	401	Inpatients and outpatients aged >18 years in a cardiovascular department
Lown et al [15]	2020	United Kingdom	Prospective, cross-sectional	Non-12-lead ECG	415	Participants aged >65 years (n=79 with AF <sup>d</sup> and n=336 without AF)
Wegner et al [16]	2020	Germany	Prospective, cross-sectional	Non-12-lead ECG	92	Inpatients with no predefined exclusion criteria
Yan et al [17]	2020	Hong Kong	Prospective, cross-sectional	PPG	44	20 patients with permanent AF and 24 control individuals in sinus rhythm
Reverberi et al [18]	2019	Italy	Prospective, longitudinal	Non-12-lead ECG	95	Patients aged >18 years diagnosed with AF and scheduled for elective cardioversion
Proesmans et al [19]	2019	Belgium	Prospective, cross-sectional	Non-12-lead ECG & PPG	223	Patients aged ≥65 years, individuals with known AF and supplemented with individuals without AF
Himmelreich et al [20]	2019	Netherlands	Prospective, cross-sectional	Non-12-lead ECG	214	Patients aged ≥18 years assigned to 12-lead ECG for any nonacute indication
Haverkamp et al [21]	2019	Norway	Prospective, cross-sectional	Non-12-lead ECG	94	Patients admitted to the cardiology ward with ongoing 12-lead ECG surveillance
Fan et al [22]	2019	China	Prospective, cross-sectional	PPG	108	Patients aged ≥18 years admitted to the hospital
Yan et al [23]	2018	Hong Kong	Prospective, cross-sectional	PPG	217	Patients admitted to the cardiology ward
William et al [24]	2018	United States	Prospective, cross-sectional	Non-12-lead ECG	52	Patients aged 35-85 years diagnosed with AF and scheduled for anti-arrhythmic drug initiation
Rozen et al [25]	2018	United States	Prospective, longitudinal	PPG	98	Patients aged >18 years diagnosed with AF and scheduled for elective DC <sup>e</sup> cardioversion
Bumgarner et al [26]	2018	United States	Prospective, longitudinal	Non-12-lead ECG	100	Patients aged 18-90 years diagnosed with AF and scheduled for elective cardioversion
Lown et al [27]	2018	United Kingdom	Prospective, cross-sectional	Non-12-lead ECG	418	Patients aged >65 years (n=82 with AF and n=336 without AF)
Desteghe et al [28]	2016	Belgium	Prospective, cross-sectional	Non-12-lead ECG	265	Patients aged ≥18 years admitted to the cardiology ward
Haberman et al [29]	2015	United States	Prospective, cross-sectional	Non-12-lead ECG	381	Athletes, students, and patients of an ambulatory cardiology clinic

<sup>a</sup>n: number of participants.

<sup>b</sup>ECG: electrocardiogram.

<sup>c</sup>PPG: photoplethysmography.

<sup>d</sup>AF: atrial fibrillation.

<sup>e</sup>DC: direct current.

### Characteristics and Quality of Studies

All the included studies followed a prospective design and had a total of 3217 participants. Of the 16 publications, 13 were

cross-sectional studies [14-17,19-24,27,28], where the measurement was conducted at a single point, and the other 3 were longitudinal [18,25,26], where the measurement was conducted more than once for each participant. The sample size



of the studies ranged from 44 to 418 participants. Four studies excluded participants aged <65 years, 4 studies recruited only participants with a diagnosis or history of AF [18,24-26], and 4 studies identified patients with AF and supplemented the cohort with control participants [15,17,19,27]. The remaining 7 studies recruited in-hospital patients with various indications [14,16,20-23,28], and 1 study recruited athletes, students, and patients of an ambulatory cardiology clinic [29]. The AF prevalence among the participants in these studies ranged from 5% to 100%.

The 16 studies investigated 2 technologies: non-12-lead ECG and PPG. This review defines non-12-lead ECG as measurements recorded by any electrode-based device; participants simply placed their fingers on the electrodes, attached the electrodes to their chest, or held the electrodes in their hands. PPG is a technology that measures changes in tissue blood volume that enables the recording of each heartbeat; the signal can be detected by any device with a camera monitoring various body parts, including the fingertip, wrist, palm, and face. Among the publications included in this review, 10 studies focused on non-12-lead ECG equipped in 10 different electrode-based ambulatory devices [15,16,18,20,21,24,26-29]; 4 studies focused on PPG recorded by cameras [17,23], phone cameras [19,22,23,25], or wristbands [14,22]; and 2 studies investigated both technologies in the same cohort [14,19]. All

measurements were processed automatically using algorithms or smartphone apps. The reference standard in all the studies was 12-lead-ECG, which was interpreted by competent physicians or cardiologists in a blinded manner. Data for the primary statistical analysis, including true positives, false positives, false negatives, and true negatives, are presented in [Multimedia Appendix 2](#).

The quality of the included studies was evaluated according to the criteria of the QUADAS-2 tool. Regarding the risk of patient selection bias, case-control studies and participants from cardiology departments were labelled as unclear risk in yellow, and studies that included only patients with AF were labelled as high risk in red. To be included in this review, all studies must have used 12-lead-ECGs as the reference standard; however, if studies used other types of reference tests in a small proportion of participants, they were still included but marked as unclear risk in the reference standard column. In the flow and timing assessment, studies that sequentially performed the reference test right before or after the index test were labelled as unclear risk, whereas studies that performed both index and reference tests simultaneously on the participants were labelled as low risk. Other criteria were assessed as provided. Two independent reviewers performed the quality appraisal, and the results are presented in [Figure 2](#).

**Figure 2.** Summary of the QUADAS-2 quality appraisal of the included studies. ECG: electrocardiogram; PPG: photoplethysmography.

Study	RISK OF BIAS				
	PATIENT SELECTION	INDEX TEST: Non-12-lead ECG	INDEX TEST: PPG	REFERENCE STANDARD	FLOW AND TIMING
Chen et al.	?	😊	😊	😊	?
Lown et al.	?	😊	-	😊	?
Wegner et al.	😊	😊	-	😊	?
Yan et al.	?	-	😊	😊	?
Reverberi et al.	😞	😊	-	😊	?
Proesmans et al.	?	😊	😊	😊	?
Himmelreich et al.	😊	😊	-	😊	😊
Haverkamp et al.	?	😊	-	😊	?
Fan et al.	?	-	😊	😊	😊
Yan et al.	😊	-	😊	😊	?
William et al.	😞	😊	-	😊	?
Rozen et al.	😞	-	😊	?	?
Bumgarner et al.	😞	😊	-	😊	😊
Lown et al.	?	😊	-	😊	?
Desteghe et al.	?	😊	-	😊	😊
Haberman et al.	😊	😊	-	😊	?

Study	APPLICABILITY CONCERN			
	PATIENT SELECTION	INDEX TEST: Non-12-lead ECG	INDEX TEST: PPG	REFERENCE STANDARD
Chen et al.	😊	😊	😊	😊
Lown et al.	😊	😊	-	😊
Wegner et al.	😊	😊	-	😊
Yan et al.	😊	-	😊	😊
Reverberi et al.	?	😊	-	😊
Proesmans et al.	😊	😊	😊	😊
Himmelreich et al.	😊	😊	-	😊
Haverkamp et al.	😊	😊	-	😊
Fan et al.	😊	-	😊	😊
Yan et al.	😊	-	😊	😊
William et al.	?	😊	-	😊
Rozen et al.	?	-	😊	😊
Bumgarner et al.	?	😊	-	😊
Lown et al.	😊	😊	-	😊
Desteghe et al.	😊	😊	-	😊
Haberman et al.	😊	😊	-	😊

😊 Low Risk    😞 High Risk    ? Unclear Risk

**Data Synthesis of Diagnostic Accuracy**

In total, 3217 participants with 7623 measurements were included in the data synthesis.

Automatic detection of AF based on non-12-lead ECG had a combined estimated sensitivity of 89.7% (95% CI 83.2%-93.9%), specificity of 95.7% (95% CI 92.0%-97.7%), PLR of 20.64 (95% CI 10.10-42.15), NLR of 0.11 (95% CI 0.06-0.19), and DOR of 224.75 (95% CI 70.10-720.56). A

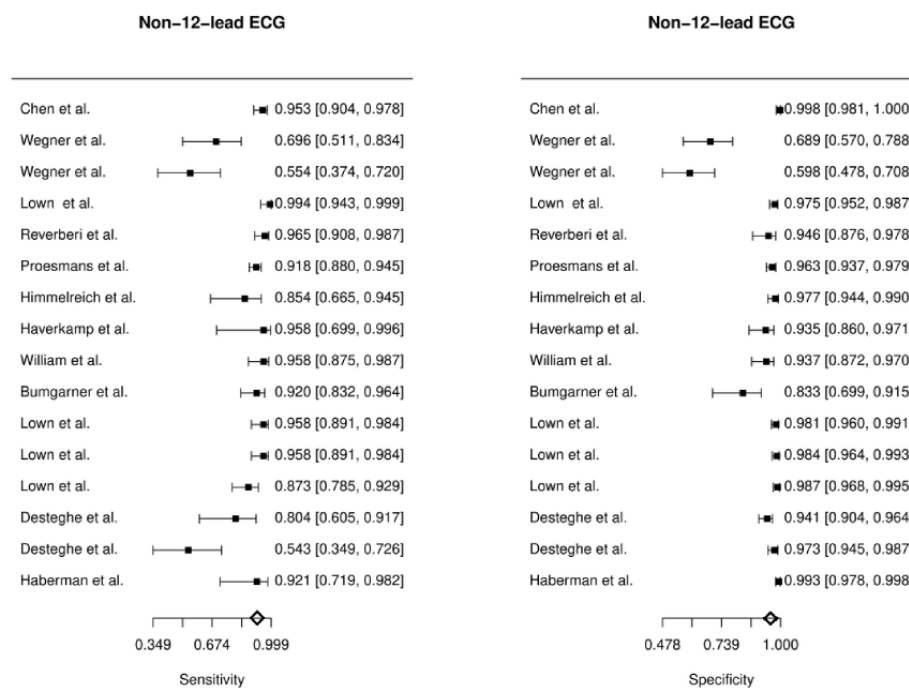
heterogeneity test among included studies was assessed with a Cochran Q of 13.99 (df=15, P=.526) and Higgins I<sup>2</sup> of 0%. Automatic detection of AF based on recordings of PPG had a combined estimated sensitivity of 94.7% (95% CI 93.3%-95.8%), specificity of 97.6% (95% CI 94.5%-99.0%), PLR of 35.51 (95% CI 18.19-69.31), NLR of 0.05 (95% CI 0.04-0.07), and DOR of 730.79 (95% CI 309.33-1726.49). A test of heterogeneity for the PPG studies reported a Cochran Q of 8.78 (df=7, P=.269) and Higgins I<sup>2</sup> of 20.26%. The forest

plots of the pooled diagnostic accuracies are presented in Figure 3 and Figure 4. SROC curves for both the technologies in all the included studies are presented in Figure 5. The AUSROCs were 0.97 for non-12-lead ECG and 0.95 for PPG.

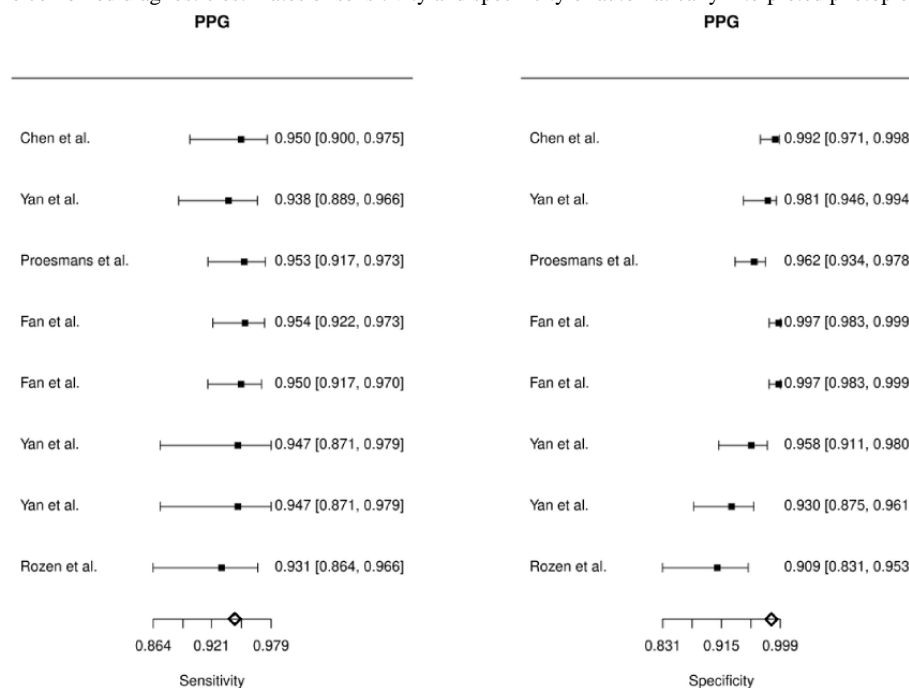
In light of the 2020 European Society of Cardiology Clinical Practice Guidelines for the diagnosis and management of AF [30] that recommended a manually interpreted, single-lead ECG  $\geq 30$  seconds as the other option to establish a definitive diagnosis of AF, we also extracted relevant data from the included studies to perform a meta-analysis of such a method.

Manual interpretation based on non-12-lead ECG had a combined sensitivity of 93.4% (95% CI 86.7%-96.8%), specificity of 96.3% (95% CI 92.9%-98.1%), PLR of 25.93 (95% CI 13.70-49.05), NLR of 0.07 (95% CI 0.04-0.14), and DOR of 439.64 (95% CI 202.89-952.65). All non-12-lead ECG segments included were recorded  $\geq 30$  seconds. The forest plots of the pooled diagnostic accuracies are shown in Multimedia Appendix 3. An SROC comparison curve between the automatic and manual interpretations of non-12-lead ECG are presented in Figure 6. No PPG recordings were examined manually among the included studies.

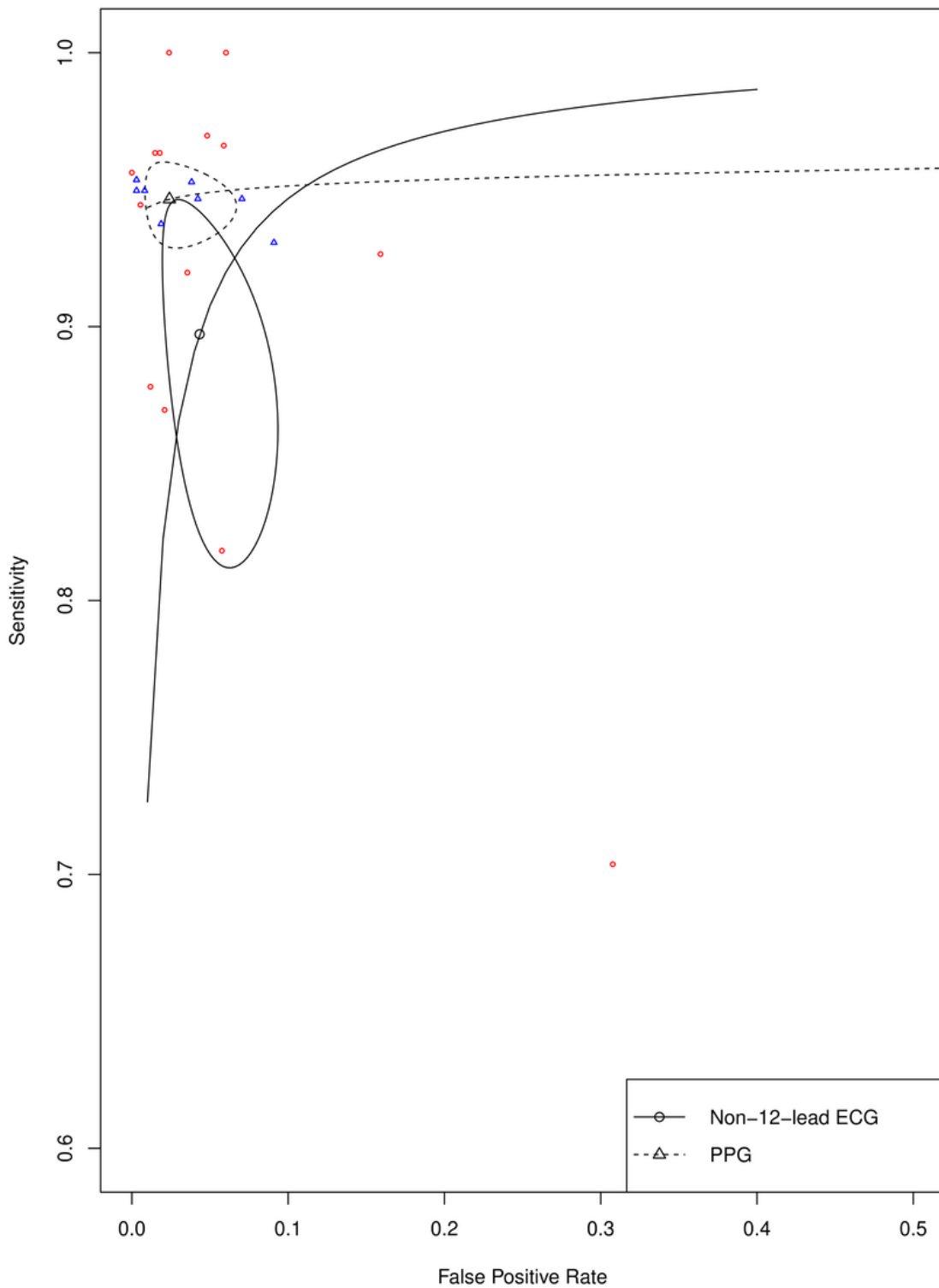
**Figure 3.** Forest plot of the combined diagnostic estimates of sensitivity and specificity of automatically interpreted non-12-lead electrocardiograms (ECGs).



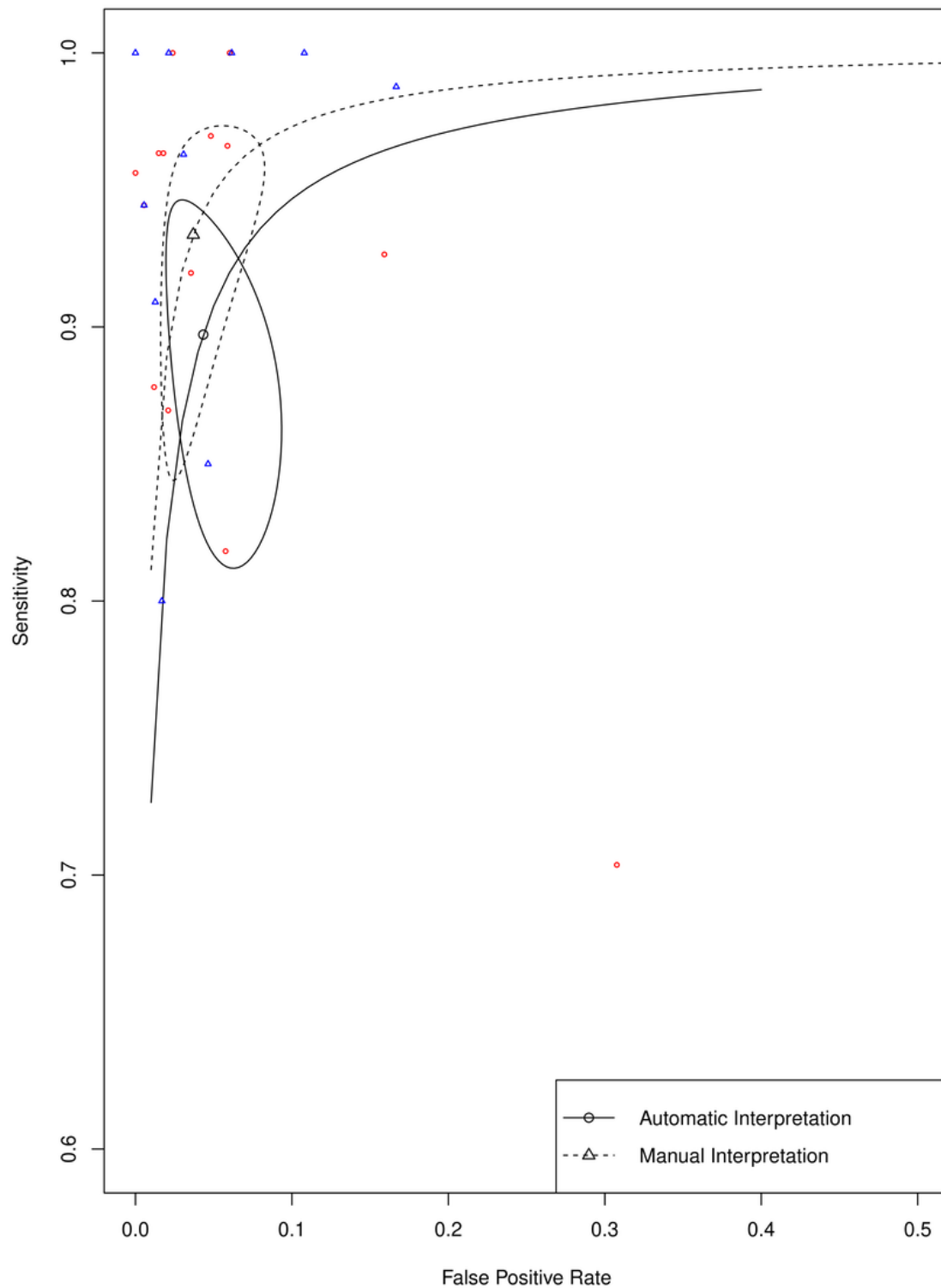
**Figure 4.** Forest plot of the combined diagnostic estimates of sensitivity and specificity of automatically interpreted photoplethymography (PPG).



**Figure 5.** Summary receiver operating curves of the automatically interpreted non-12-lead electrocardiogram (ECG) and photoplethymography (PPG) in the diagnosis of atrial fibrillation.



**Figure 6.** Summary receiver operating curves of automatic and manual interpretations of non-12-lead electrocardiogram in the diagnosis of atrial fibrillation.

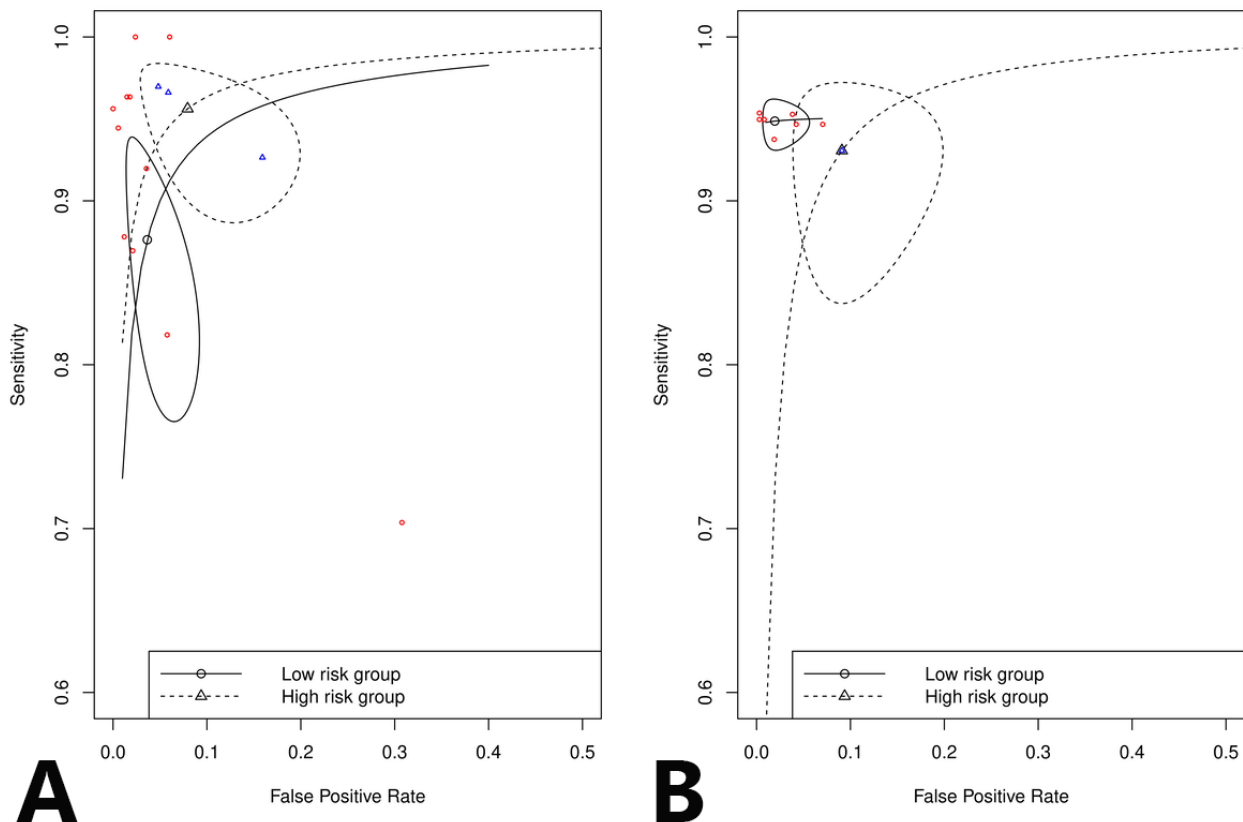


**Subgroup Analysis: Study Population**

A subgroup analysis (Figure 7) was performed to investigate the effect of the study design and population on the diagnostic

accuracy. Studies were divided into 2 groups: low risk (group 1), including those that recruited participants with and without AF, and high risk (group 2), including those that only recruited participants with prediagnosed AF.

**Figure 7.** Subgroup analysis of the study population, including a comparison of summary receiver operating curves between the low-risk study population, which included patients with and without atrial fibrillation, and high-risk study population, which included only patients with atrial fibrillation, for (A) non-12-lead electrocardiogram and (B) photoplethysmography.



Non-12-lead ECG yielded a sensitivity of 87.6% and specificity of 96.4% in low-risk studies ( $n=13$ ) and a sensitivity of 95.6% and specificity of 92.1% in high-risk studies ( $n=3$ ). PPG yielded a sensitivity of 94.9% and specificity of 98.1% in low-risk studies ( $n=7$ ) and a sensitivity of 93.1% and specificity of 90.9% in high-risk studies ( $n=1$ ). The subgroup analysis of non-12-lead ECG indicated a higher sensitivity and lower specificity among the high-risk studies. On the other hand, only 1 study was included in the high-risk group for PPG; thus, no conclusion could be made regarding the subgroup analysis of the PPG technology.

### Fagan Nomogram

The Fagan nomogram analysis (Multimedia Appendix 4 and Multimedia Appendix 5) demonstrated that, with an age-adjusted prevalence of 0.5% in the general population with pooled PLR of 20.64 and pooled NLR of 0.11, the posttest probabilities of non-12-lead ECG increased to 9.40% and decreased to 0.06%, respectively; with pooled PLR of 35.51 and pooled NLR of 0.05, the posttest probabilities of PPG increased to 15.14% and decreased to 0.03%, respectively. By contrast, in a high-risk population with a prevalence of 2.3%, the application of non-12-lead ECG had posttest probabilities of 32.70% and 0.26%, respectively, and PPG had posttest probabilities of 45.53% and 0.12%, respectively.

### Publication Bias

Publication bias was assessed using Deek funnel plots (Multimedia Appendix 6 and Multimedia Appendix 7). For studies investigating the non-12-lead ECG method, visual

inspection of the funnel plot indicated a likely absence of publication bias ( $P=.107$ ). For studies investigating the PPG method, a visual evaluation of the funnel plot also provided no clear evidence of publication bias ( $P=.107$ ).

## Discussion

### Principal Findings

This systematic review and meta-analysis on the diagnostic accuracy of 16 studies published between 2015 and 2021 confirmed that both non-12-lead ECG and PPG are highly accurate technologies in detecting AF. Automatically interpreted PPG provided the highest sensitivity and specificity for the diagnosis of AF, immediately followed by the manual interpretation of non-12-lead ECG, whereas automatically diagnosed AF based on non-12-lead ECG performed relatively weaker than the first 2 methods. However, they all demonstrated outstanding diagnostic accuracy. The differences among these approaches may be attributed to measuring techniques: PPG optically records volumetric changes in local arterioles and, in this manner, measures pulse variability (or R-R interval) and heart rhythm variability, while non-12-lead ECG assesses the electrical activity of the heart using electrodes. Other possible explanations for the variability include different classification thresholds and different methods used in the algorithm design.

Early diagnosis of AF is crucial for preventing AF-related morbidity, mortality, and economic burden. The prevalence of undiagnosed AF in the United States is estimated to be 1%-2%, and the incremental cost burden could amount to US \$3.1 billion

per year [31]. In comparison, smart wearable devices and mobile phones are almost ubiquitous worldwide, and the incremental cost for their application is relatively low. Our summarized findings suggest that the appropriate application of these technologies enables early diagnosis and treatment, which might possibly contribute to reducing the cost of illness. The current guidance suggests pulse palpation and 12-lead ECG in AF screening. However, in a meta-analysis by Taggar et al [6], pulse palpation was proved inferior to 4 other methods because of lower specificity; additionally, the use of 12-lead ECG is limited by its inconvenience and the fact that it captures an ECG at only 1 time point. In clinical practice, continuous Holter monitoring is one of the most commonly used methods to detect AF. However, this examination requires patients to carry the machine with them for an extended period, which results in inconvenience; thus, it is also not a suitable tool for general screening.

The 2 main technologies reviewed in this meta-analysis, non-12-lead ECG and PPG, have several advantages as screening tools for AF. First, they involve lightweight, low-cost, and ambulatory devices. In this review, non-12-lead ECGs were employed in devices such as handheld tablets [16,19-21,24,27,28], watchbands [14,26], handheld rod-like sensors [28], and electrodes attached to the chest [15,16,18,19,27]. By contrast, PPG signals can be assessed on the fingertips [19,22,23,25], wrist [14,22], earlobe, and face [17,23] by any device with simple optoelectronic components, such as a smartphone. Second, these technologies require a relatively short monitoring time, which enables fast and timely screening; the monitoring time most commonly varied from 30 seconds to 1 minute. Third, measurements obtained by these devices were automatically interpreted by built-in algorithms, which means that the tests can be conducted without the presence of a health care professional. Several studies [16,20,24,26,28] included in this review conducted additional analyses comparing algorithms' and physicians' interpretations of the same non-12-lead ECG segments and concluded that automated algorithm performance is not inferior to competent professional interpretation. Finally, as demonstrated in the results, the automatically generated diagnoses established with both technologies yielded outstanding diagnostic accuracy. Automatically interpreted PPG had a sensitivity and specificity of 94.7% and 97.6%, respectively. Automatically interpreted non-12-lead ECG had a sensitivity and specificity of 89.7% and 95.7%, respectively. As for the interpretation of non-12-lead ECG by competent physicians, an established method to diagnose AF according to the 2020 European Society of Cardiology Clinical Guidelines had a sensitivity and specificity of 93.4% and 96.3%, respectively. Demonstrating effectiveness, convenience, and time savings with high diagnostic accuracies, we suggest using these technologies with built-in automatic interpretation as preliminary screening tools for the detection of AF when the gold standard method, which generally requires a physician's interpretation, is not feasible. Notably, the screening program should target high-risk populations (eg, elderly) to avoid false positives.

We performed a subgroup analysis of both technologies on account of the apparent patient selection bias among included

studies. The sensitivity and specificity of a test are generally believed to not vary with the disease prevalence of a population; however, variations of these diagnostic parameters were often spotted. According to a previous article that summarized the phenomenon and proposed several possible causes [32], the specificity of a test tended to be lower with high disease prevalence, and although not significant, the sensitivity appeared higher with high prevalence of some diseases. This review also observed instability of the sensitivity and specificity between low-risk groups (studies that recruited patients with and without AF) and high-risk groups (studies that recruited only patients with AF). Higher sensitivity and lower specificity were generated in the high-risk groups using the non-12-lead ECG tests; however, no conclusion could be made for the PPG method because only 1 study included a high-risk group. The results indicate that the performance of these technologies was affected by the recruited population and design of the included studies. Future validation conducted in a more general population is warranted to investigate such an occurrence.

### Strengths and Limitations

To our knowledge, this is the first systematic review and meta-analysis to compare the diagnostic accuracy of 2 common technologies, namely non-12-lead ECG and PPG, used in ambulatory devices for detecting AF. This review followed PRISMA guidelines, implemented a comprehensive search strategy, applied strict inclusion and exclusion criteria, and employed 2 independent reviewers to assess all the included studies. For instance, one of the criteria was to exclude studies that failed to provide a timely reference standard; this resulted in the exclusion of multiple large-scale screening studies but ensured time consistency between the index test and reference standard. Moreover, we investigated the heterogeneity and publication bias of included studies, as well as the posttest probability of AF using these ambulatory device technologies. Finally, an additional subgroup analysis was performed to investigate the effect of the study population on diagnostic performance. The result identified sensitivity and specificity variations between low-risk and high-risk populations, indicating future validation of the diagnostic accuracy of these tests is needed in a more general population.

This review of studies on AF detection using ambulatory devices has several limitations. The most noteworthy concern is the study population investigated. Except for the 4 studies that recruited only patients with AF for their assessment, other studies were mostly conducted in a case-control style or recruited inpatients from hospitals. This problem was reflected in the QUADAS-2 quality assessment and possibly contributed to the instability of diagnostic accuracy in the subgroup analysis. Additionally, the heterogeneity of devices and algorithms used should be considered. Although we explored the accuracy of ambulatory devices from a technology perspective, these technologies were applied in diverse devices, and the measurements were automatically interpreted by their respective algorithms. Furthermore, measurements of insufficient quality or unclassified by algorithms tended to be excluded in the calculation of the diagnostic accuracy in some of the studies [14,19,22,24,26]; the proportion of insufficient or unclassified recordings ranged from 0.5% to 33.8%. Finally, some studies

regarded atrial flutter as the same disease state as AF, viewing the incident as a positive AF result [16,19,20,24,26-29].

### Conclusions

Both non-12-lead ECG and PPG technologies offered high diagnostic accuracies for AF. Automatically interpreted PPG recordings generated the highest sensitivity and specificity compared to both the manual and automatic interpretations of

non-12-lead ECG. Detection of AF employing automatic analysis techniques may serve as a useful preliminary screening tool before administering a gold standard test, which generally requires analyses by competent physicians. Subgroup analysis indicated variations of sensitivity and specificity between studies that recruited low-risk and high-risk populations, and future validation of these diagnostic tests in the general population is warranted.

### Acknowledgments

The authors thank Mark Lown, Felix K Wegner, and Aviram Hochstadt for sharing their original data and contributing to the analytic calculations. The authors also thank Wallace Academic Editing for manuscript editing. We also thank the Higher Education Sprout Project of the Ministry of Education (MOE) in Taiwan for financially supporting this work [DP2-110-21121-01-A-04].

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

The literature search strategy for the systematic review and meta-analysis in Medline (Ovid) and Embase databases.

[DOCX File, 14 KB - [mhealth\\_v9i4e26167\\_app1.docx](#)]

#### Multimedia Appendix 2

Details of the study population, ambulatory devices, measuring time, reference standard, prevalence, and measurement data of the included studies.

[DOCX File, 17 KB - [mhealth\\_v9i4e26167\\_app2.docx](#)]

#### Multimedia Appendix 3

Forest plot of the combined diagnostic estimates of sensitivity and specificity of manually interpreted non-12-lead electrocardiogram.

[PNG File, 83 KB - [mhealth\\_v9i4e26167\\_app3.png](#)]

#### Multimedia Appendix 4

Fagan nomogram for the evaluation of the clinical utility of automatically interpreted non-12-lead electrocardiogram (ECG).

[DOCX File, 196 KB - [mhealth\\_v9i4e26167\\_app4.docx](#)]

#### Multimedia Appendix 5

Fagan nomogram for the evaluation of the clinical utility of automatically interpreted photoplethysmography (PPG).

[DOCX File, 189 KB - [mhealth\\_v9i4e26167\\_app5.docx](#)]

#### Multimedia Appendix 6

Deek funnel plots for the assessment of publication bias of non-12-lead electrocardiogram (ECG) studies.

[PDF File (Adobe PDF File), 101 KB - [mhealth\\_v9i4e26167\\_app6.pdf](#)]

#### Multimedia Appendix 7

Deek funnel plots for the assessment of publication bias of photoplethysmography (PPG) studies.

[PDF File (Adobe PDF File), 154 KB - [mhealth\\_v9i4e26167\\_app7.pdf](#)]

### References

1. Patel NJ, Atti V, Mitrani RD, Viles-Gonzalez JF, Goldberger JJ. Global rising trends of atrial fibrillation: a major public health concern. *Heart* 2018 Dec;104(24):1989-1990. [doi: [10.1136/heartjnl-2018-313350](#)] [Medline: [29907645](#)]
2. Odotayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016 Sep 06;354:i4482 [FREE Full text] [doi: [10.1136/bmj.i4482](#)] [Medline: [27599725](#)]
3. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, ESC Scientific Document Group. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016 Oct 07;37(38):2893-2962. [doi: [10.1093/eurheartj/ehw210](#)] [Medline: [27567408](#)]



4. Jones N, Taylor C, Hobbs F, Bowman L, Casadei B. Screening for atrial fibrillation: a call for evidence. *Eur Heart J* 2020 Mar 07;41(10):1075-1085 [FREE Full text] [doi: [10.1093/eurheartj/ehz834](https://doi.org/10.1093/eurheartj/ehz834)] [Medline: [31811716](https://pubmed.ncbi.nlm.nih.gov/31811716/)]
5. National Institute for Health and Care Excellence. Atrial fibrillation: management (Clinical guideline [CG180]). National Institute for Health and Care Excellence. URL: <https://www.nice.org.uk/guidance/cg180> [accessed 2021-01-25]
6. Taggar JS, Coleman T, Lewis S, Heneghan C, Jones M. Accuracy of methods for detecting an irregular pulse and suspected atrial fibrillation: A systematic review and meta-analysis. *Eur J Prev Cardiol* 2016 Aug;23(12):1330-1338 [FREE Full text] [doi: [10.1177/2047487315611347](https://doi.org/10.1177/2047487315611347)] [Medline: [26464292](https://pubmed.ncbi.nlm.nih.gov/26464292/)]
7. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. *N Engl J Med* 2019 Nov 14;381(20):1909-1917. [doi: [10.1056/nejmoa1901183](https://doi.org/10.1056/nejmoa1901183)]
8. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. In: *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. The Cochrane Collaboration. London, United Kingdom: The Cochrane Collaboration, 2009; Oct 27, 2009.
9. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011 Oct 18;155(8):529-536 [FREE Full text] [doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009)] [Medline: [22007046](https://pubmed.ncbi.nlm.nih.gov/22007046/)]
10. Philipp D, Heinz H. Meta-Analysis of Diagnostic Accuracy with mada. *cran.r-project.org*. URL: <https://cran.r-project.org/web/packages/mada/vignettes/madapdf.%202012> [accessed 2021-01-25]
11. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005 Sep;58(9):882-893. [doi: [10.1016/j.jclinepi.2005.01.016](https://doi.org/10.1016/j.jclinepi.2005.01.016)] [Medline: [16085191](https://pubmed.ncbi.nlm.nih.gov/16085191/)]
12. 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](https://doi.org/10.1161/CIRCULATIONAHA.113.005119)] [Medline: [24345399](https://pubmed.ncbi.nlm.nih.gov/24345399/)]
13. Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. *Thromb Haemost* 2017 Dec 04;110(08):213-222. [doi: [10.1160/th13-02-0165](https://doi.org/10.1160/th13-02-0165)]
14. Chen E, Jiang J, Su R, Gao M, Zhu S, Zhou J, et al. A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. *Heart Rhythm* 2020 May;17(5 Pt B):847-853 [FREE Full text] [doi: [10.1016/j.hrthm.2020.01.034](https://doi.org/10.1016/j.hrthm.2020.01.034)] [Medline: [32354449](https://pubmed.ncbi.nlm.nih.gov/32354449/)]
15. Lown M, Brown M, Brown C, Yue AM, Shah BN, Corbett SJ, et al. Machine learning detection of Atrial Fibrillation using wearable technology. *PLoS One* 2020 Jan 24;15(1):e0227401 [FREE Full text] [doi: [10.1371/journal.pone.0227401](https://doi.org/10.1371/journal.pone.0227401)] [Medline: [31978173](https://pubmed.ncbi.nlm.nih.gov/31978173/)]
16. Wegner FK, Kochhäuser S, Ellermann C, Lange PS, Frommeyer G, Leitz P, et al. Prospective blinded Evaluation of the smartphone-based AliveCor Kardia ECG monitor for Atrial Fibrillation detection: The PEAK-AF study. *Eur J Intern Med* 2020 Mar;73:72-75. [doi: [10.1016/j.ejim.2019.11.018](https://doi.org/10.1016/j.ejim.2019.11.018)] [Medline: [31806411](https://pubmed.ncbi.nlm.nih.gov/31806411/)]
17. Yan BP, Lai WHS, Chan CKY, Au ACK, Freedman B, Poh YC, et al. High-Throughput, Contact-Free Detection of Atrial Fibrillation From Video With Deep Learning. *JAMA Cardiol* 2020 Jan 01;5(1):105-107 [FREE Full text] [doi: [10.1001/jamacardio.2019.4004](https://doi.org/10.1001/jamacardio.2019.4004)] [Medline: [31774461](https://pubmed.ncbi.nlm.nih.gov/31774461/)]
18. Reverberi C, Rabia G, De Rosa F, Bosi D, Botti A, Benatti G. The RITMIA™ Smartphone App for Automated Detection of Atrial Fibrillation: Accuracy in Consecutive Patients Undergoing Elective Electrical Cardioversion. *Biomed Res Int* 2019;2019:4861951 [FREE Full text] [doi: [10.1155/2019/4861951](https://doi.org/10.1155/2019/4861951)] [Medline: [31355264](https://pubmed.ncbi.nlm.nih.gov/31355264/)]
19. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile Phone-Based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: Diagnostic Accuracy Study of the FibrCheck App. *JMIR Mhealth Uhealth* 2019 Mar 27;7(3):e12284 [FREE Full text] [doi: [10.2196/12284](https://doi.org/10.2196/12284)] [Medline: [30916656](https://pubmed.ncbi.nlm.nih.gov/30916656/)]
20. Himmelreich JC, Karregat EP, Lucassen WA, van Weert HC, de Groot JR, Handoko ML, et al. Diagnostic Accuracy of a Smartphone-Operated, Single-Lead Electrocardiography Device for Detection of Rhythm and Conduction Abnormalities in Primary Care. *Ann Fam Med* 2019 Sep 09;17(5):403-411 [FREE Full text] [doi: [10.1370/afm.2438](https://doi.org/10.1370/afm.2438)] [Medline: [31501201](https://pubmed.ncbi.nlm.nih.gov/31501201/)]
21. Haverkamp HT, Fosse SO, Schuster P. Accuracy and usability of single-lead ECG from smartphones - A clinical study. *Indian Pacing Electrophysiol J* 2019;19(4):145-149 [FREE Full text] [doi: [10.1016/j.ipej.2019.02.006](https://doi.org/10.1016/j.ipej.2019.02.006)] [Medline: [30794928](https://pubmed.ncbi.nlm.nih.gov/30794928/)]
22. Fan Y, Li Y, Li J, Cheng W, Shan Z, Wang Y, et al. Diagnostic Performance of a Smart Device With Photoplethysmography Technology for Atrial Fibrillation Detection: Pilot Study (Pre-mAFA II Registry). *JMIR Mhealth Uhealth* 2019 Mar 05;7(3):e11437 [FREE Full text] [doi: [10.2196/11437](https://doi.org/10.2196/11437)] [Medline: [30835243](https://pubmed.ncbi.nlm.nih.gov/30835243/)]
23. Yan BP, Lai WHS, Chan CKY, Chan SC, Chan L, Lam K, et al. Contact-Free Screening of Atrial Fibrillation by a Smartphone Using Facial Pulsatile Photoplethysmographic Signals. *J Am Heart Assoc* 2018 Apr 05;7(8):1 [FREE Full text] [doi: [10.1161/JAHA.118.008585](https://doi.org/10.1161/JAHA.118.008585)] [Medline: [29622592](https://pubmed.ncbi.nlm.nih.gov/29622592/)]
24. William AD, Kanbour M, Callahan T, Bhargava M, Varma N, Rickard J, et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: The iREAD Study. *Heart Rhythm* 2018 Oct;15(10):1561-1565. [doi: [10.1016/j.hrthm.2018.06.037](https://doi.org/10.1016/j.hrthm.2018.06.037)] [Medline: [30143448](https://pubmed.ncbi.nlm.nih.gov/30143448/)]

25. Rozen G, Vaid J, Hosseini SM, Kaadan MI, Rafael A, Roka A, et al. Diagnostic Accuracy of a Novel Mobile Phone Application for the Detection and Monitoring of Atrial Fibrillation. *Am J Cardiol* 2018 May 15;121(10):1187-1191. [doi: [10.1016/j.amjcard.2018.01.035](https://doi.org/10.1016/j.amjcard.2018.01.035)] [Medline: [29525063](https://pubmed.ncbi.nlm.nih.gov/29525063/)]
26. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch Algorithm for Automated Detection of Atrial Fibrillation. *J Am Coll Cardiol* 2018 May 29;71(21):2381-2388 [FREE Full text] [doi: [10.1016/j.jacc.2018.03.003](https://doi.org/10.1016/j.jacc.2018.03.003)] [Medline: [29535065](https://pubmed.ncbi.nlm.nih.gov/29535065/)]
27. Lown M, Yue AM, Shah BN, Corbett SJ, Lewith G, Stuart B, et al. Screening for Atrial Fibrillation Using Economical and Accurate Technology (From the SAFETY Study). *Am J Cardiol* 2018 Oct 15;122(8):1339-1344 [FREE Full text] [doi: [10.1016/j.amjcard.2018.07.003](https://doi.org/10.1016/j.amjcard.2018.07.003)] [Medline: [30131106](https://pubmed.ncbi.nlm.nih.gov/30131106/)]
28. Desteghe L, Raymaekers Z, Lutin M, Vijgen J, Dilling-Boer D, Koopman P, et al. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. *Europace* 2017 Jan;19(1):29-39. [doi: [10.1093/europace/euw025](https://doi.org/10.1093/europace/euw025)] [Medline: [26893496](https://pubmed.ncbi.nlm.nih.gov/26893496/)]
29. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN, et al. Wireless Smartphone ECG Enables Large-Scale Screening in Diverse Populations. *J Cardiovasc Electrophysiol* 2015 May 19;26(5):520-526. [doi: [10.1111/jce.12634](https://doi.org/10.1111/jce.12634)] [Medline: [25651872](https://pubmed.ncbi.nlm.nih.gov/25651872/)]
30. Hindricks G, Potpara T, Dagres N. Corrigendum to: 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* 2021 Feb 01;42(5):507. [doi: [10.1093/eurheartj/ehaa798](https://doi.org/10.1093/eurheartj/ehaa798)] [Medline: [33029625](https://pubmed.ncbi.nlm.nih.gov/33029625/)]
31. Turakhia M, Shafrin J, Bogner K, Trocio J, Abdulsattar Y, Wiederkehr D, et al. Economic Burden of Undiagnosed Nonvalvular Atrial Fibrillation in the United States. *Journal of the American College of Cardiology* 2015 Mar;65(10):A475. [doi: [10.1016/s0735-1097\(15\)60475-2](https://doi.org/10.1016/s0735-1097(15)60475-2)]
32. Leeftang MMG, Rutjes AWS, Reitsma JB, Hooft L, Bossuyt PMM. Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ* 2013 Aug 06;185(11):E537-E544 [FREE Full text] [doi: [10.1503/cmaj.121286](https://doi.org/10.1503/cmaj.121286)] [Medline: [23798453](https://pubmed.ncbi.nlm.nih.gov/23798453/)]

## Abbreviations

**AF:** atrial fibrillation

**AUSROC:** area under the summary receiver operating curve

**DOR:** diagnostic odds ratio

**ECG:** electrocardiogram

**NLR:** negative likelihood ratio

**PLR:** positive likelihood ratio

**PPG:** photoplethysmography

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

**QUADAS:** quality assessment of diagnostic accuracy studies

**SROC:** summary receiver operating characteristic

*Edited by L Buis; submitted 02.12.20; peer-reviewed by P Vandervoort, S Lewis, B Vaes; comments to author 22.01.21; revised version received 07.02.21; accepted 11.03.21; published 09.04.21.*

*Please cite as:*

Yang TY, Huang L, Malwade S, Hsu CY, Chen YC

Diagnostic Accuracy of Ambulatory Devices in Detecting Atrial Fibrillation: Systematic Review and Meta-analysis

*JMIR Mhealth Uhealth* 2021;9(4):e26167

URL: <https://mhealth.jmir.org/2021/4/e26167>

doi: [10.2196/26167](https://doi.org/10.2196/26167)

PMID: [33835039](https://pubmed.ncbi.nlm.nih.gov/33835039/)

©Tien Yun Yang, Li Huang, Shwetambara Malwade, Chien-Yi Hsu, Yang Ching Chen. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 09.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

---

Review

# Biofeedback-Based Connected Mental Health Interventions for Anxiety: Systematic Literature Review

---

Mahra Alneyadi<sup>1</sup>, BSc; Nidal Drissi<sup>1</sup>, MSc; Mariam Almeqbaali<sup>1</sup>, BSc; Sofia Ouhbi<sup>1</sup>, PhD

Department of Computer Science & Software Engineering, College of Information Technology, Al Ain, Abu Dhabi, United Arab Emirates

---

**Corresponding Author:**

Sofia Ouhbi, PhD

Department of Computer Science & Software Engineering

College of Information Technology

United Arab Emirates University

Al Ain, Abu Dhabi, 15551

United Arab Emirates

Phone: 971 3 713 ext 5568

Email: [sofia.ouhbi@uaeu.ac.ae](mailto:sofia.ouhbi@uaeu.ac.ae)

## Abstract

---

**Background:** Connected mental health, which refers to the use of technology for mental health care and technology-based therapeutic solutions, has become an established field of research. Biofeedback is one of the approaches used in connected mental health solutions, which is mainly based on the analysis of physiological indicators for the assessment and management of the psychological state. Biofeedback is recommended by many therapists and has been used for conditions including depression, insomnia, and anxiety. Anxiety is associated with several physiological symptoms, including muscle tension and breathing issues, which makes the inclusion of biofeedback useful for anxiety detection and management.

**Objective:** The aim of this study was to identify interventions using biofeedback as a part of their process for anxiety management and investigate their perceived effectiveness.

**Methods:** A systematic literature review of publications presenting empirically evaluated biofeedback-based interventions for anxiety was conducted. The systematic literature review was based on publications retrieved from IEEE Digital Library, PubMed, ScienceDirect, and Scopus. A preliminary selection of papers was identified, examined, and filtered to include only relevant publications. Studies in the final selection were classified and analyzed to extract the modalities of use of biofeedback in the identified interventions, the types of physiological data that were collected and analyzed and the sensors used to collect them. Processes and outcomes of the empirical evaluations were also extracted.

**Results:** After final selection, 13 publications presenting different interventions were investigated. The interventions addressed either primarily anxiety disorders or anxiety associated with health issues such as migraine, Parkinson disease, and rheumatology. Solutions combined biofeedback with other techniques including virtual reality, music therapy, games, and relaxation practices and used different sensors including cardiovascular belts, wrist sensors, or stretch sensors to collect physiological data such as heart rate, respiration indicators, and movement information. The interventions targeted different cohorts including children, students, and patients. Overall, outcomes from the empirical evaluations yielded positive results and emphasized the effectiveness of connected mental health solutions using biofeedback for anxiety; however, certain unfavorable outcomes, such as interventions not having an effect on anxiety and patients' preferring traditional therapy, were reported in studies addressing patients with specific physical health issues.

**Conclusions:** The use of biofeedback in connected mental health interventions for the treatment and management of anxiety allows better screening and understanding of both psychological and physiological patient information, as well as of the association between the two. The inclusion of biofeedback could improve the outcome of interventions and boost their effectiveness; however, when used with patients suffering from certain physical health issues, suitability investigations are needed.

(*JMIR Mhealth Uhealth* 2021;9(4):e26038) doi:[10.2196/26038](https://doi.org/10.2196/26038)

---

**KEYWORDS**

anxiety; biofeedback; systematic literature review; mental health; eHealth; mHealth; connected health; digital health

## Introduction

### Background

Anxiety is the brain's way of reacting to stress and alerting to possible danger, which makes it an expected feeling in a person's daily life, as it can be triggered by normal daily scenarios [1,2]. However, the persistence of this feeling might be an indicator that the person is suffering from an anxiety disorder. *Anxiety disorder* is an umbrella term that comprises various mental disorders characterized by excessive anxiety, tension, and fear that interfere with the person's daily life and interrupt the normal execution of daily tasks and activities [2]. There are 5 main types of anxiety disorders [3]: generalized anxiety disorder, obsessive compulsive disorder, panic disorder, posttraumatic stress disorder, and phobia. Most anxiety disorders affect women 2 times more than they affect men [3]. The causes of anxiety disorders are not clear and differ from one person to another based on many factors. Anxiety can be triggered by difficult life experiences, surrounding environment, and health behavior or can be caused by physical factors, such as overactive brain areas involved in emotions and behavior, genetics, and brain chemistry [4,5]. In addition to psychological symptoms, anxiety is also associated with many physiological symptoms including muscle tension, heartbeat issues, breathing issues, sweating, dry mouth, and headaches [1,3,5].

Connected mental health is the subfield of connected health that refers to the use of information and communication technologies for mental health care and includes all related areas such as mobile mental health, digital mental health, tele-mental health and e-mental health [6]. Connected mental health now plays a crucial role in the health care sector and contributes to improving the delivery of mental health care by providing novel, affordable, and easy-to-access solutions [6,7]. Technology has been included in mental health care in several forms including the exploitation of sensors for mental states' detection and management [8], mobile apps for mental health care [9], and websites [10,11]. Use of technology has also facilitated access to many therapeutic solutions for anxiety disorders and mental health generally [7]. Examples include virtual reality (VR), computer- and internet-based cognitive behavioral therapy, and biofeedback [7]. VR is a 3D environment that can be either similar to or different from the real world. It allows the patients to interact with a specific environment based on the psychological issue and feared stimuli [12]. Computer-based cognitive behavioral therapy and self-rated mental health help evaluate the patient's mental health via apps or websites by analyzing the user's answers to certain mental state assessment questions [13]. Patients can also use chatbots and online therapy services to communicate with mental health professionals and obtain a diagnosis [13,14]. Biofeedback-based mental health interventions aim to identify the patient's mental state by monitoring body activities [15].

Biofeedback for mental health is based on measuring physiological changes associated with psychological states [16] to help monitor the body functions that are affected by the psychological reactions. Biofeedback-based interventions use and monitor different physiological factors, including heart rate,

galvanic skin response (also known as electrodermal response), and respiration measurements. The main aim of biofeedback training is to provide patients with awareness and insight on their physiological changes, helping them better control those changes, and consequently, better control their mental state. Biofeedback has been shown to be one of the useful ways to help reduce the symptoms of anxiety disorders [17]. The physiological manifestations of anxiety make biofeedback useful in anxiety detection and treatment solutions.

Biofeedback could be useful for several mental health issues, such as stress, anxiety, hypertension, and depression [18]. Moreover, advances in technology have allowed biofeedback to become affordable, cost-effective, and easily used by practitioners as well as users [18]. The aim of this paper was to investigate the use of biofeedback in connected mental health solutions for anxiety disorders by conducting a systematic literature review. We investigated modalities of biofeedback use by identifying treatment approaches combined with it, types of sensors used in the interventions, and the physiological data collected and analyzed. In addition, we reviewed empirical evidence on the effectiveness of biofeedback-based interventions for anxiety from intervention outcomes.

### Related Work

Biofeedback is becoming one of the complementary and alternative medicine forms recommended by many doctors and therapists [16]. This section presents examples of literature addressing the adoption of biofeedback in treatment solutions for anxiety disorders and other mental issues.

Biofeedback in combination with psychotherapy was used for military medical providers suffering from anxiety, depression, and insomnia [19]. Psychotherapy helped reduce symptoms of anxiety and depression but could not improve insomnia issues. Yet when combined with biofeedback, the treatment was able to improve the sleep of the military medical providers [19].

A portable biofeedback device was integrated into clinical practice for patients with anxiety who were receiving cognitive behavioral therapy-based treatment [20]. Patients reported higher satisfaction with biofeedback-based treatment compared to that reported for other relaxation techniques such as meditation, yoga, and unassisted breathing. It was reported that biofeedback could be a promising treatment adjunct for disorders of autonomic arousal and could be easily integrated into treatment [20].

An exploratory review [21] investigated the efficacy of treating anxiety disorders in children and adolescents using biofeedback, cognitive behavioral therapy, and mindfulness combined with technological tools and programs, including serious games, web-based tools, apps, and internet-based tools. It reported that connected health interventions were found to be effective for anxiety management, as many studies reported they were as effective as traditional treatments.

Moreover, other studies [22-24] discussed the effectiveness of biofeedback in serious games for emotion regulation and mental health, generally. They reported that there was promising evidence for integrating biofeedback in serious games for managing anxiety [22-24]. Biofeedback has been reported as a

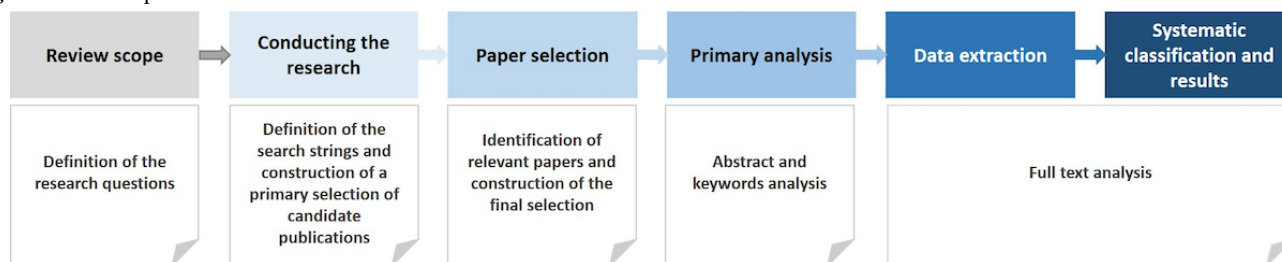
successful approach to practice emotion regulation and was also found to improve performance on decision-making tasks in serious games [22]. Moreover, examples of existing biofeedback-based games have shown positive results for reducing depression and anxiety [24].

## Methods

### Overview

We aimed to investigate the modalities of use of biofeedback in connected mental health solutions for anxiety disorders, as well as investigate the empirical evidence on such solutions. This study follows the quality reporting guidelines set out by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis [25]). Figure 1 summarizes the review process.

Figure 1. Review process.



### Research Questions

We attempted to answer the following 2 research questions: (1) What are the biofeedback-based connected mental health interventions for anxiety available in the literature? (2) How effective are biofeedback-based connected mental health solutions in treating anxiety? The first aims to identify the different treatment approaches that could be combined with biofeedback for the management of anxiety, as well as the different types of sensors and physiological information that could be used in biofeedback-based interventions. The second

aims to identify patients’ interaction with the biofeedback treatments in connected mental health solutions and how beneficial it is to their case and level of anxiety.

### Research Method

The search for candidate papers was conducted in IEEE Digital Library, PubMed, ScienceDirect, and Scopus using the search strings presented in Table 1. The search strings included relevant terms and were formulated to identify a wide selection of candidate publications.

Table 1. Search strings.

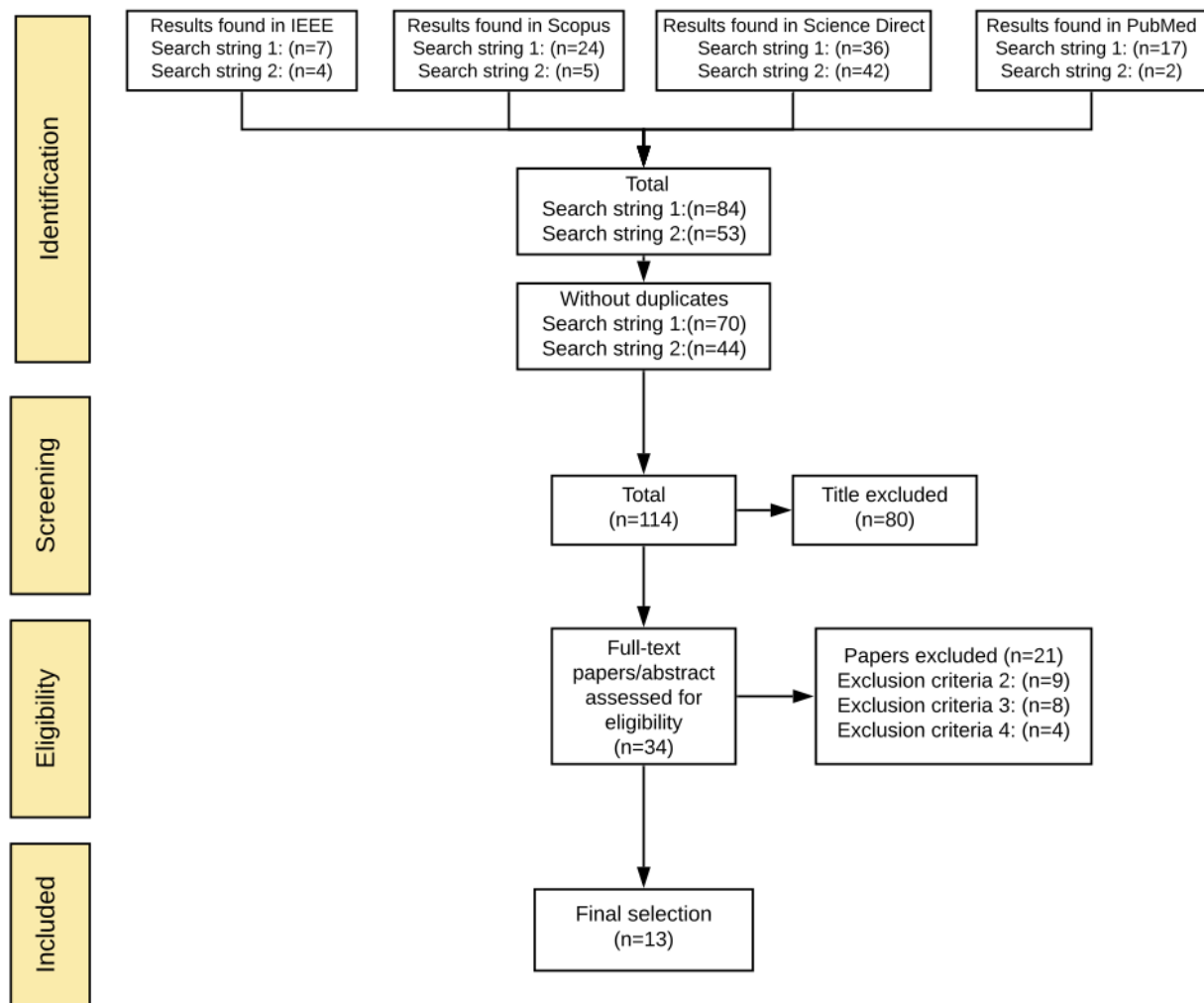
ID	Search string
1	“biofeedback” AND “treatment” AND “anxiety” AND “disorder” AND (“e-health” OR “m-health” OR “digital health” OR “ehealth” OR “mhealth” OR “connected health” OR “technology”)
2	(“computer” OR “mobile”) AND “biofeedback” AND “anxiety” AND “disorder” AND “games”

### Paper Selection

After the primary selection of candidate papers by applying the search strings to the digital libraries, duplicates were removed and the candidate publications were first analyzed by inspecting the titles then filtered based on a set of eligibility criteria to construct a final selection of relevant studies (Figure 2). Papers were selected based on the following inclusion criteria: studies that (1) addressed treating anxiety using biofeedback, and (2)

combined information and communication technologies and biofeedback to treat anxiety. Exclusion criteria were any of the following: (1) publications that were not original research papers (eg, index or abstract); (2) studies that presented biofeedback as a treatment for general mental health or other issues, such as depression, but not anxiety; (3) studies that did not combine information and communication technologies and biofeedback to treat anxiety; or (4) studies that were not empirically evaluated (eg, reviews).

Figure 2. Selection process.



## Data Collection

Studies in the final selection were analyzed to extract relevant information. A data extraction form with predefined fields was developed in a spreadsheet (Microsoft Excel 2016). Data were extracted and categorized. While extracting data, we focused mainly on answering the research questions by identifying treatment approaches combined with biofeedback for anxiety management, the cohort groups addressed by the interventions, types of physiological data analyzed in the interventions, sensors used to collect the data, the processes of the empirical evaluations conducted on the interventions, and the outcome of empirical evaluations. Studies were then classified, mainly by grouping the interventions by cohort group and treatment approach.

## Synthesis

The synthesis method used in this study consisted of reading and analyzing selected studies, categorizing the data extracted from selected studies, and classifying the studies by enumerating the number of interventions for each data category. It should be noted that interventions including more than one treatment

approach, more than one type of physiological data, or more than one type of sensor were counted in each category. The results are presented in figures and tables, because visualization of the results facilitated their analysis, and as a narrative summary describing the interventions and principal findings.

## Results

### Search Results

A total of 114 candidate publications were identified; however, only 13 studies met the eligibility criteria and were included in the final selection (Figure 2); 10 were published from 2011 to 2020, 2 were published in 2009, and 1 was published in 2002. The selected studies targeted people from different cohort groups: 8 of the interventions targeted mainly anxiety, while the rest addressed health issues associated with anxiety. In interventions that were identified, biofeedback was combined with different treatment techniques. Information on the publications, including targeted health issues and targeted cohort groups, are presented in Table 2, and information on processes and outcomes of the empirical evaluations are presented in Multimedia Appendix 1.

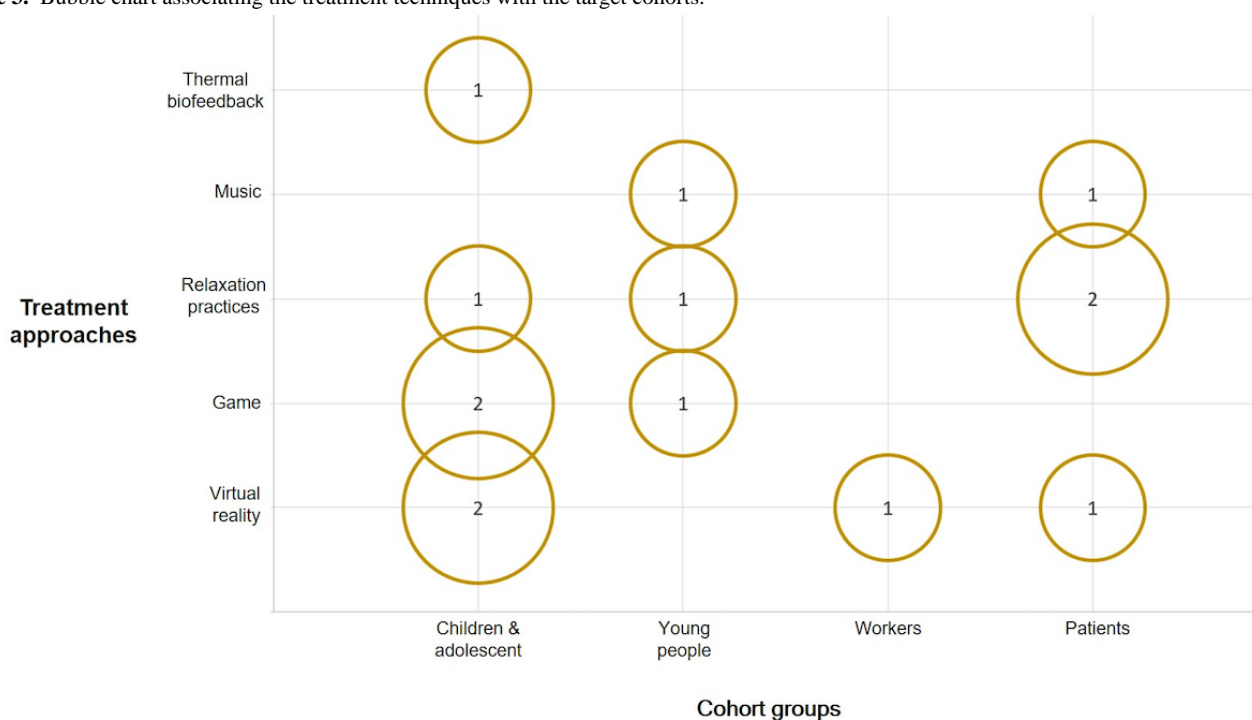
**Table 2.** Final selection.

Age category/reference	Year of publication	Cohort group	Condition
<b>Adults</b>			
[26]	2020	Patients with Parkinson disease	Anxiety
[27]	2019	Patients with rheumatic diseases	Pain and anxiety
[28]	2018	Patients with drug-resistant temporal lobe epilepsy seizures	Associated anxiety, stress, and depression
[29]	2014	Teachers and nurses	Psychological stress including anxiety
[30]	2009	Patients with generalized anxiety disorder	Anxiety
<b>Young adults</b>			
[31]	2018	Youth in residential care	Anxiety
[32]	2016	College students	Anxiety
[33]	2009	University students	Anxiety related to their studies
<b>Children and adolescents</b>			
[34]	2017	Children and adolescents with autism	Anxiety and performance
[35]	2016	Children	Risk of anxiety
[36]	2015	Pediatric patients	Pain and anxiety
[24]	2011	Children and adolescents	Anxiety
[37]	2002	Children	Migraine and anxiety

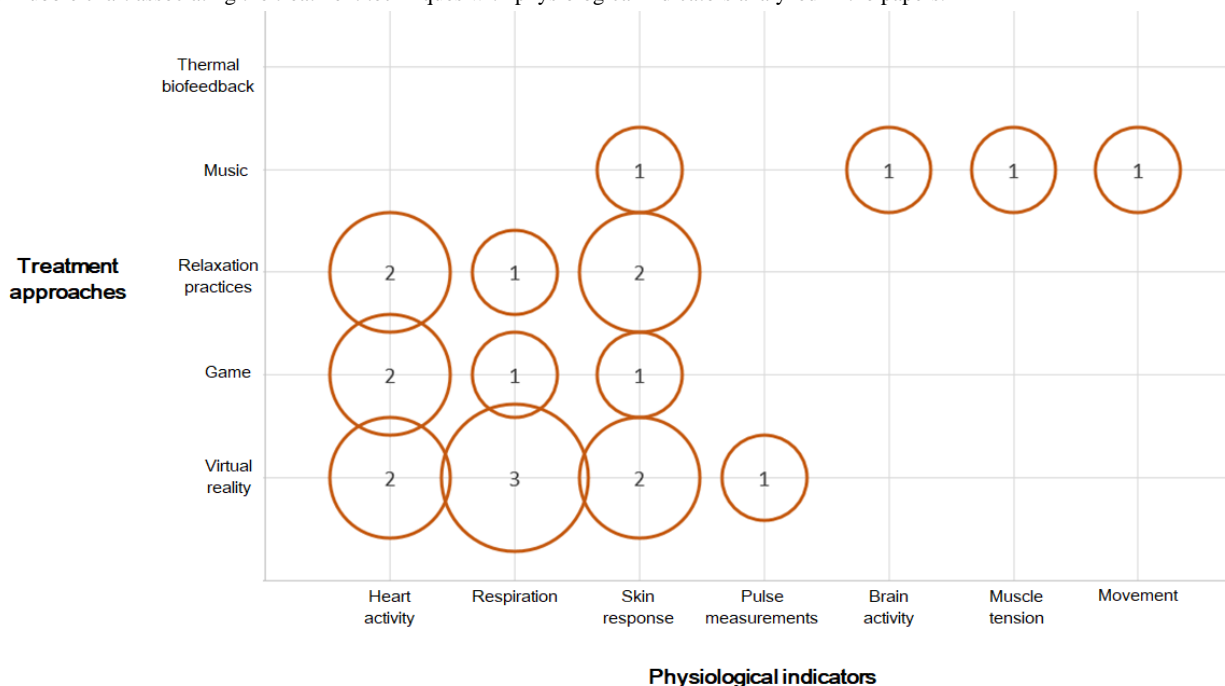
Identified treatment techniques included biofeedback with VR, games, and relaxation practices used with different cohort groups including children and adolescents, young adults, and patients (Figure 3). Children and adolescents were the age group most

targeted by interventions. Figure 4 presents the association between treatment techniques and physiological data types. Data related to heart activity and skin response were used the most.

**Figure 3.** Bubble chart associating the treatment techniques with the target cohorts.



**Figure 4.** Bubble chart associating the treatment techniques with physiological indicators analyzed in the papers.



### Interventions for Adults

Of the 13 studies, 5 presented solutions for adults [26-30], 4 of which addressed patients [26-28,30]. Two of the solutions for patients were based on VR [27,30]. One was a VR-based biofeedback and guided meditation system for the management of pain and anxiety in patients with rheumatic disease [27] consisting of modules with and without respiratory biofeedback, monitored using a microphone. The user, when interacting with the VR environment, is instructed to breathe along to an oscillating pacer. Patients who used the intervention preferred the guided meditation module that did not include biofeedback over the module with biofeedback [27]. The other VR-based solution [30] was for the treatment of generalized anxiety disorder using VR and mobile phones, in which the user explored a virtual island by following a therapist-recorded narrative that included relaxation exercises. The user was also connected to biosensors that record physiological parameters including skin conductance, heart rate, and respiration. Some of the virtual experience elements were directly modified by the real-time heart rate variability of the user. Including biofeedback in the VR-based mobile intervention resulted in a higher decrease of anxiety than the VR-based mobile intervention without biofeedback [30].

Music therapy was also used with biofeedback; one intervention [26] consisted of an Ambuloso system for patients with Parkinson disease. The system is based on a method of musical gait training that rewards desirable gait behaviors with music play. The system comprised wireless headphones and a combination of a music player and a wearable sensor (iOS gyroscope and accelerometer) worn above the knee of the user to collect movement and step information. During gait biofeedback training, music is played when the user achieves a set stride length. Even though the data used in this intervention were not related to psychological status, the intervention had a

secondary impact on patients' levels of depression and anxiety [26].

A biofeedback system based on skin conductance used for the management of nonpharmacological epilepsy and associated anxiety, depression, and stress was proposed in [28]. The system presents the users with movies to trigger certain emotions that users were required to control [28]. While the users watched the films, skin conductance was measured and recorded using electrodes on the index and middle finger of the left hand of the participant. The study reported that skin conductance response was not related to changes in anxiety levels, which is contradictory to the findings of other studies in which skin conductance was used as an indicator of anxiety [38,39]. Yet this contradiction might be specific to the group targeted by the intervention. Even though the intervention reported a decrease in patients' anxiety and depression, it was suggested that this decrease reflects a nonspecific or a placebo effect of using the intervention [28].

For adult workers, particularly teachers and nurses [29], an intervention consisting of virtual scenarios with real-time monitoring was used for the management of the psychological state. The system put the user in virtual scenarios of stressful experiences and natural scenarios to learn certain relaxation techniques while the user was connected to cardiovascular belt and wrist biosensors to collect heart rate and heart rate variability. Data collected by the sensors were assessed by a decision support system that provided users with a real-time graphical representation of their current stress level. Some of the elements of the virtual experience were driven by the emotional status of the patient, measured by the biosensors. In this intervention, the combination of VR with biofeedback showed significantly better results than those of VR interventions without biofeedback or traditional treatments. The combination of VR and biofeedback showed better results in



reducing anxiety when compared to traditional cognitive behavioral therapy treatment [29].

### Interventions for Young People

Of the 13 studies, 3 addressed mainly young people [31-33]. One of the interventions was a biofeedback video game called Dojo for anxiety and externalizing problems for young people in residential care [31]. The game promoted emotion regulation by providing tutorials on cognitive behavioral therapy-based relaxation techniques such as deep breathing, progressive muscle relaxation, positive thinking, and guided imagery. The game also included mini games that trigger different emotions such as fear, frustration, and anger and require the users to regulate their emotions using the techniques. User heart rate was monitored while playing the game through a biofeedback hardware and was displayed on the screen. Users were required to control their physiological reaction to succeed in the game, thus encouraging them to regulate their emotions [31].

Two interventions [32,33] targeted college students. One was based on music therapy and biofeedback treatment for anxiety [32]. The intervention provided the user with a relaxing music experience and judged the effect by measuring dynamic changes including skin resistance, as well as electroencephalography and electromyography indicators [32]. Another intervention provided biofeedback training to help university students overcome anxiety related to their studies [33]. The intervention is based on using a stress sweeper biofeedback device for the physiological measurements. Collected physiological indicators included heart rate and respiration. The biofeedback device helped guide participants while they practiced techniques to reduce anxiety related to their studies [33].

### Interventions for Children and Adolescents

Of the 13 studies, 5 presented solutions for children and adolescents [24,34-37]: 3 studies presented interventions based on games [24,35,36], including the programs *Freeze-Framer* and *Journey to the Wild Divine*, to reduce anxiety and depression. The programs were biofeedback-assisted relaxation training games that included cognitive behavioral therapy. The interventions used 2 to 3 electrodes to record momentary changes in heart rate variability and skin conductance levels. An example of a biofeedback relaxation activity offered by the programs is one in which users built a bridge—the user was required to relax to continue the construction of the bridge. As user breathing slows and tension decreases, the bridge is built; if the user experiences frustration or anxiety, the bridge disappears. After a continuous period of relaxation, the bridge is finished and the user can cross it [24]. *Journey to the Wild Divine* [36] collects heart rate variability and skin temperature of the users via sensors attached to the users' fingers. The program contains 15 levels, and completing one level requires meeting certain relaxation standards. Quick completion of a level is an indication of the ability of the patient to quickly regulate their physiological response and relax [36].

A VR biofeedback breathing game called DEEP [35] situates users in an underwater fantasy world, where they can move

freely. The game is not based on levels or goals; rather, it provides personal breathing and meditation support by promoting diaphragmatic breathing through biofeedback. The game collects breathing data using a stretch sensor. A microcontroller interprets the sensor readings and sends data to the game, where it is used to change the user experience and manipulate the game elements. Slow and deep breathing allows players to move forward in the game, thus promoting diaphragmatic breathing. The intervention showed promising results; it was reported to decrease anxiety in a timely manner [35].

VR was also used in an anxiety-sensitive and performance-sensitive adaptive system for children with autism [34]. The system was composed of different modules including VR-based social communication, real-time physiological data acquisition, intelligent anxiety predictor, and strategy generator modules. The system captured photoplethysmography, electrodermal activity, and skin temperature signals to use as anxiety predictors. The system's purpose was to identify and quantify the user's anxiety level from real-time biomarkers, along with performance metrics during social communication simulations. The system adapts, progressing through different levels of tasks, based on the anxiety and performance metrics [34].

A system providing thermal biofeedback treatment to children suffering from migraine [37] included two types of thermal treatments, hand-warming biofeedback and hand-cooling biofeedback, to help the children deal with their migraine symptoms and anxiety. The treatments had different effects on migraines, yet did not affect anxiety levels [37].

### Empirical Evaluations

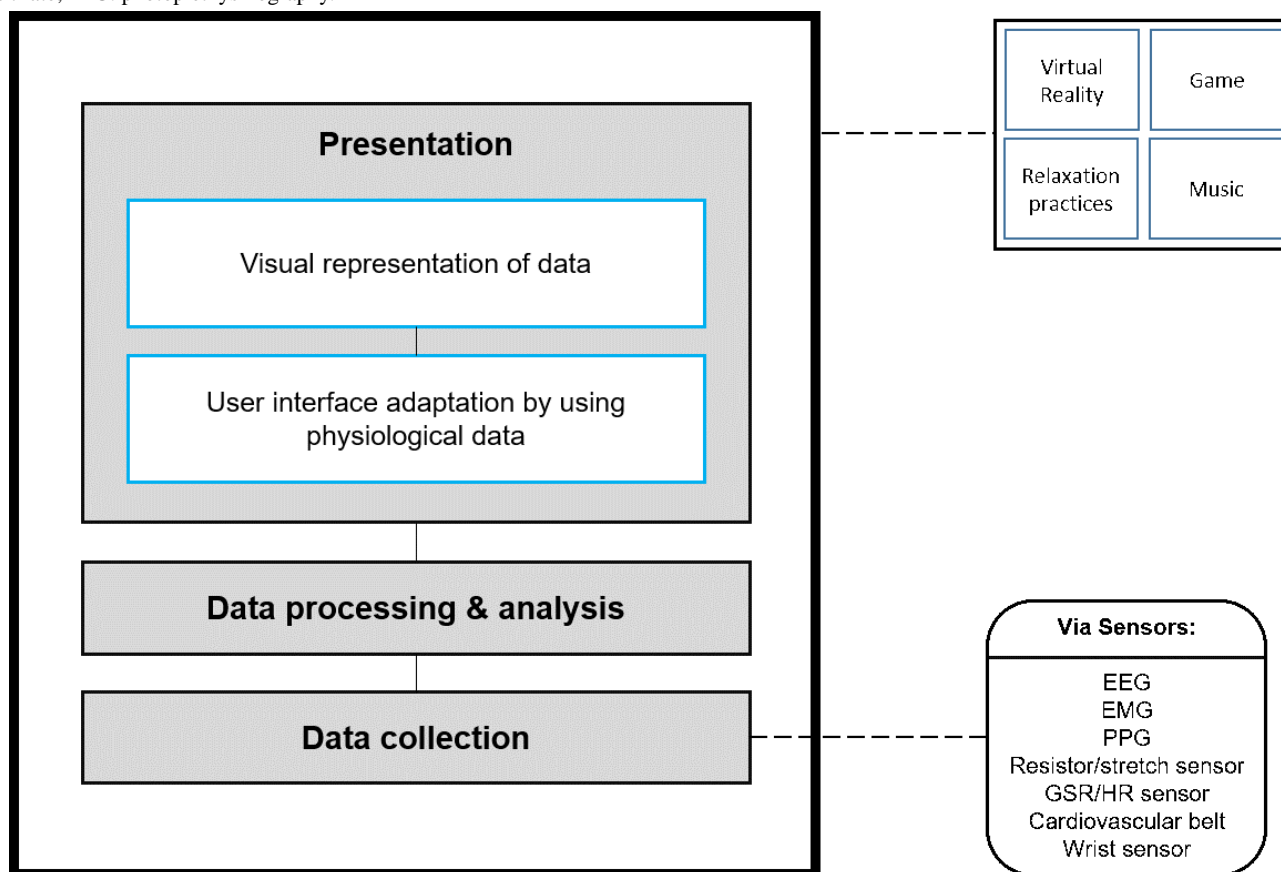
All 13 studies presented processes and outcomes of empirical evaluations. The empirical evaluations, in general, followed similar protocols: they included participants based on sets of eligibility criteria and divided participants into intervention and control groups who were wait-listed or received traditional treatment. Overall, the empirical evaluations yielded positive results. However, unfavorable outcomes were reported in some studies (Multimedia Appendix 1).

## Discussion

### Main Findings

Interventions identified in this review used biofeedback with different treatment techniques, different sensors and physiological data, and for different purposes; however, all the interventions followed the same logic (Figure 5). The process begins with physiological data collection via sensors attached to the user. Physiological data are then processed and analyzed to generate a type of feedback from the system. In the interventions identified in this review, 2 main types of feedback were reported. Based on the readings of the physiological data the interventions either modified the user interface and experience or offered a visual presentation of the physiological changes.

**Figure 5.** Layer diagram for a biofeedback-based system. EEG: electroencephalography; EMG: electromyography; GSR: galvanic skin response; HR: heart rate; PPG: photoplethysmography.



**Biofeedback With VR and Games**

Biofeedback was mostly found to be combined with VR and was found to increase the efficacy of such interventions [29,30]. The VR concept emerged in the 1980s and is based on transforming any real situation into a virtual experience. VR is based on the basic elements of generating images, presenting sensory information, and updating the displayed images based on the users’ position and orientation [40]. VR has the potential to create rich sensory experiences [41] and has been used for the management of different mental issues, such as the treatment of anxiety [42]. VR has been widely used for anxiety, by exposing patients to virtual situations triggering anxiety and teaching them how to deal with such situations. VR-based treatments have been used for different anxiety issues [43], including phobia management [44], mainly for the delivery of exposure therapy. VR facilitates exposure therapy exercises because they can be conducted in the therapist’s office rather than in real-world phobic situations [44]. VR exposure has also been found to be effective for panic disorder [45] and in helping veterans with posttraumatic stress disorder [46].

Having issues interacting with the world is at the core of many psychological issues, including anxiety. VR could be a useful approach to face that issue, as it helps to simulate real-life experiences for the patients. Examples include helping manage phobia issues, which are characterized by intense fear and anxiety when interacting with certain elements, and posttraumatic stress disorder, in which patients have flashbacks of traumatizing experiences [42]. VR and biofeedback were

also used with children [35], suggesting that it can be a children-friendly treatment. VR was used with children with multiple disabilities to familiarize them with the use of the wheelchair [47]. Moreover, VR has also been adopted into different types of games for purposes such as interactive entertainment, interactive training, and education [48]. The combination of VR with games has also been used in different medical settings as well, including medical training [49] and rehabilitation treatments [50].

Games have been coupled with biofeedback for mental health. Games for health can be classified as serious games, which mainly refers to games that do not focus on enjoyment, entertainment, and fun as their main purpose [51] but rather on elements such as education, training, and health improvement [52]. Quality games have been shown to influence behavior [53] and enhance concentration [54], as well as to facilitate learning and information retention [55,56]. Blending biofeedback with games has shown promising results. A computer biofeedback game based on animated gut imagery was used with people with irritable bowel syndrome and helped patients decrease their stress as well as mitigate symptoms related to their physical disorder [57]. Biofeedback-based games have also been used for physical issues management, for example, in balance training for people affected by chronic hemiplegic stroke, which showed positive results and was reported to be a feasible adjunct to conventional therapy [58].

Interventions including games mainly targeted the young generation [24,31] and have been shown to be effective. Young

people are heavy users of technology, in general, and are the most familiar with gaming technologies, as millions of adolescents and young adults play video games and are the primary users of games [59,60]. Heavy use of technology and familiarity with games makes the young generation a suitable group for the use of game-based mental health care interventions. When introduced to VR interventions and games, biofeedback helped enhance the virtual and gaming experiences and encouraged user engagement, in addition to providing insight on user physiological and psychological states to help with the management of health issues. However, it must be noted that for such interventions to be helpful, user needs, experience, and preferences should be at the heart of their design [23,42].

### **Biofeedback With Relaxation Practices and Music**

Relaxation techniques and practices include somatic methods such as progressive relaxation, breathing, stretching, and physical exercises, as well as cognitive approaches including imagery, meditation, goal-directed visualization, and self-awareness [61]. These methods have been shown to be effective in mitigating anxiety and stress [61,62]. Biofeedback was coupled with such practices in some interventions [33,36] to inform users about the changes in their physiological measurements to help them better understand the association between their psychological and physiological indicators. Biofeedback inclusion was effective in the reduction of anxiety in these interventions [33,36].

Representing results with graphs, diagrams, animations, or other in order to deliver information for better understanding is used in education and is viewed as an effective teaching tool [63]. The same approach was used with patients to educate them about their health, which resulted in the patients being more satisfied, becoming more knowledgeable, and gaining understanding about their health issues [64]. In the case of anxiety, helping patients better understand and follow their case, the physiological changes associated with their mental issue, and treatment modalities can help them be more aware of the effect of anxiety on their bodies and help them be more accepting and trusting of the treatment applied, which might even influence the outcome of the treatment. Relaxation techniques were also used for pediatric pain management [33]. Relaxation techniques may be effective complementary therapies to pharmacologic techniques for pediatric pain, which can help reduce or even eliminate the amount of medication needed to treat the pain [65]. Relaxation approaches have also been used to manage other types of pain, including labor pain during childbirth [66], perioperative pain [67], and chronic pain [68].

Music with biofeedback was also among the relaxation techniques that showed positive results [32]. Music therapy has been shown to regulate both the physical and the mental health by affecting both the physiology and psychology of the person. Music can be combined with and included in different mental health management approaches and relaxation techniques [69] and can be used as a relaxation method on its own, as it was reported to be as effective as the progressive muscle relaxation method [70]. Because preferred or accepted types of music can be different in each community, when including music in

psychological treatments, cultural consideration might be necessary [69].

Relaxation practices and music-based techniques have been widely adopted in technology-based interventions for anxiety, mainly in mobile apps [14]. Relaxation practices and music influence not only the psychology but also the physiology of our bodies. The inclusion of biofeedback could help improve mental care interventions based on those techniques. Biofeedback provides insight on physiological changes, which helps patients both assess and manage their psychological state.

### **Biofeedback for Anxiety Associated With Other Health Issues**

Anxiety was found to be common among patients of many physical disorders including patients with cancer [71] and people suffering from chronic obstructive pulmonary disease [72]. Anxiety is also common when undergoing medical procedures, such as magnetic resonance imaging [73]. Therefore, some interventions addressed symptoms of specific health issues and the anxiety associated with their prevalence [27,28,34]. Those interventions included solutions based on combining biofeedback with VR, which, in certain cases, showed unfavorable results [27]. This might be due to VR not being suitable for patients with health issues such as migraine, headache, seizure disorder, and vestibular abnormalities [43]. In addition, anxiety caused by the prevalence of specific health issues is generally impacted by the progress of the health issue itself, which might indicate that, in the interventions addressing other health issues, anxiety levels were impacted as a result of the change in the user's health but not as a direct effect of the intervention on anxiety specifically [26,37].

### **Implications**

This review may be of interest to researchers, mental health interventions' developers, and practitioners interested in the use of biofeedback in mental health management as it presents descriptions and analysis of 13 examples of biofeedback-based interventions for anxiety and different treatment approaches that could be combined with biofeedback, including VR, games, and music therapy. We also presented physiological indicators that could be exploited in biofeedback-based interventions and analyzed for anxiety management including heart rate measurements, respiration, and movement, which could be collected using different sensors including wrist sensors, cardiovascular belts, electroencephalography, and electromyography. The review showed different cohorts that could benefit from biofeedback-based interventions including children, patients, and workers. The interventions have generally yielded positive results in improving anxiety.

Use of biofeedback allowed better screening, understanding, and control of physiological factors during anxiety management interventions, improving outcomes and effectiveness; however, a need for additional investigation for certain health issues was highlighted.

It must be noted that this review might have some limitations: (1) the inclusion of other terms in the search string might result in additional publications and (2) searching Google Scholar might have resulted other relevant studies.

For future work, we intend to use the findings of this review to collaborate with mental health care professionals to create a biofeedback-based mobile app for the management of anxiety for young adults in United Arab Emirates.

## Acknowledgments

This work is part of the Abu Dhabi Young Investigator Award 2019 (AYIA19-001) awarded by the Abu Dhabi Research and Development Authority and the Startup project (31T131) funded by the United Arab Emirates University.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Empirical evaluations.

[DOCX File, 28 KB - [mhealth\\_v9i4e26038\\_app1.docx](#)]

## References

1. Anxiety disorders. National Institute of Mental Health. 2018. URL: <https://www.nimh.nih.gov/health/topics/anxiety-disorders/index.shtml> [accessed 2021-01-04]
2. Anxiety disorders. WebMD. 2020. URL: <https://www.webmd.com/anxiety-panic/guide/anxiety-disorders#1> [accessed 2021-01-04]
3. Nutt DJ. Overview of diagnosis and drug treatments of anxiety disorders. *CNS Spectr* 2005 Jan 07;10(1):49-56. [doi: [10.1017/s1092852900009901](https://doi.org/10.1017/s1092852900009901)] [Medline: [15618947](#)]
4. Martin EI, Ressler KJ, Binder E, Nemeroff CB. The neurobiology of anxiety disorders: brain imaging, genetics, and psychoneuroendocrinology. *Psychiatr Clin North Am* 2009 Sep;32(3):549-575 [FREE Full text] [doi: [10.1016/j.psc.2009.05.004](https://doi.org/10.1016/j.psc.2009.05.004)] [Medline: [19716990](#)]
5. Generalised anxiety disorder in adults. National Health Service. 2018. URL: <https://www.nhs.uk/conditions/generalised-anxiety-disorder/> [accessed 2021-01-04]
6. Drissi N, Ouhbi S, Janati Idrissi MA, Fernandez-Luque L, Ghogho M. Connected mental health: systematic mapping study. *J Med Internet Res* 2020 Aug 28;22(8):e19950 [FREE Full text] [doi: [10.2196/19950](https://doi.org/10.2196/19950)] [Medline: [32857055](#)]
7. Przeworski A, Newman MG. Efficacy and utility of computer-assisted cognitive behavioural therapy for anxiety disorders. *Clin Psychol* 2006 Aug 20;10(2):43-53. [doi: [10.1080/13284200500378779](https://doi.org/10.1080/13284200500378779)]
8. Drissi N, Ouhbi S, García-Berná J, Idrissi M, Ghogho M. Sensor-based solutions for mental health care: a systematic literature review. In: Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies. 2020 Presented at: 13th International Conference on Health Informatics; February 24-26; Valletta, Malta p. 593-600. [doi: [10.5220/0009100805930600](https://doi.org/10.5220/0009100805930600)]
9. Drissi N, Ouhbi S, Idtissi M, Ghogho M. Gamification-based apps for PTSD: an analysis of functionality and characteristics. Abu Dhabi, United Arab Emirates; 2019 Presented at: 16th ACS/IEEE International Conference on Computer Systems and Applications; November 3-7; Abu Dhabi, United Arab Emirates. [doi: [10.1109/aiccsa47632.2019.9035325](https://doi.org/10.1109/aiccsa47632.2019.9035325)]
10. Zermatten A, Khazaal Y, Coquard O, Chatton A, Bondolfi G. Quality of web-based information on depression. *Depress Anxiety* 2010 Sep;27(9):852-858. [doi: [10.1002/da.20665](https://doi.org/10.1002/da.20665)] [Medline: [20099271](#)]
11. Barnes C, Harvey R, Wilde A, Hadzi-Pavlovic D, Wilhelm K, Mitchell PB. Review of the quality of information on bipolar disorder on the internet. *Aust N Z J Psychiatry* 2009 Jan;43(10):934-945. [doi: [10.1080/00048670903179137](https://doi.org/10.1080/00048670903179137)]
12. Maples-Keller JL, Bunnell BE, Kim S, Rothbaum BO. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harv Rev Psychiatry* 2017;25(3):103-113 [FREE Full text] [doi: [10.1097/HRP.000000000000138](https://doi.org/10.1097/HRP.000000000000138)] [Medline: [28475502](#)]
13. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. *PLoS One* 2015;10(3):e0119895 [FREE Full text] [doi: [10.1371/journal.pone.0119895](https://doi.org/10.1371/journal.pone.0119895)] [Medline: [25786025](#)]
14. Drissi N, Ouhbi S, Janati Idrissi MA, Ghogho M. An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps. *Int J Med Inform* 2020 Sep 30;141:104243 [FREE Full text] [doi: [10.1016/j.ijmedinf.2020.104243](https://doi.org/10.1016/j.ijmedinf.2020.104243)] [Medline: [32768994](#)]
15. Ratanasiripong P, Ratanasiripong N, Kathalae D. Biofeedback intervention for stress and anxiety among nursing students: a randomized controlled trial. *ISRN Nurs* 2012;2012:827972 [FREE Full text] [doi: [10.5402/2012/827972](https://doi.org/10.5402/2012/827972)] [Medline: [22811932](#)]
16. Frank DL, Khorshid L, Kiffer JF, Moravec CS, McKee MG. Biofeedback in medicine: who, when, why and how? *Ment Health Fam Med* 2010 Jun;7(2):85-91 [FREE Full text] [Medline: [22477926](#)]

17. Ratanasiripong P, Kaewboonchoo O, Ratanasiripong N, Hanklang S, Chumchai P. Biofeedback intervention for stress, anxiety, and depression among graduate students in public health nursing. *Nurs Res Pract* 2015;2015:160746 [FREE Full text] [doi: [10.1155/2015/160746](https://doi.org/10.1155/2015/160746)] [Medline: [25954515](https://pubmed.ncbi.nlm.nih.gov/25954515/)]
18. Ratanasiripong P, Sverduk K, Prince J, Hayashino D. Biofeedback and counseling for stress and anxiety among college students. *J Coll Stud Dev* 2012;53(5):742-749. [doi: [10.1353/csd.2012.0070](https://doi.org/10.1353/csd.2012.0070)]
19. McLay RN, Spira JL. Use of a portable biofeedback device to improve insomnia in a combat zone, a case report. *Appl Psychophysiol Biofeedback* 2009 Dec 5;34(4):319-321. [doi: [10.1007/s10484-009-9104-3](https://doi.org/10.1007/s10484-009-9104-3)] [Medline: [19655243](https://pubmed.ncbi.nlm.nih.gov/19655243/)]
20. Reiner R. Integrating a portable biofeedback device into clinical practice for patients with anxiety disorders: results of a pilot study. *Appl Psychophysiol Biofeedback* 2008 Mar;33(1):55-61. [doi: [10.1007/s10484-007-9046-6](https://doi.org/10.1007/s10484-007-9046-6)] [Medline: [18286369](https://pubmed.ncbi.nlm.nih.gov/18286369/)]
21. Tozzi F, Nicolaidou I, Galani A, Antoniadis A. eHealth interventions for anxiety management targeting young children and adolescents: exploratory review. *JMIR Pediatr Parent* 2018 May 10;1(1):e5 [FREE Full text] [doi: [10.2196/pediatrics.7248](https://doi.org/10.2196/pediatrics.7248)] [Medline: [31518330](https://pubmed.ncbi.nlm.nih.gov/31518330/)]
22. Jerčić P, Sundstedt V. Practicing emotion-regulation through biofeedback on the decision-making performance in the context of serious games: a systematic review. *Entertain Comput* 2019 Mar;29:75-86. [doi: [10.1016/j.entcom.2019.01.001](https://doi.org/10.1016/j.entcom.2019.01.001)]
23. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. *Front Psychiatry* 2016;7:215. [doi: [10.3389/fpsy.2016.00215](https://doi.org/10.3389/fpsy.2016.00215)] [Medline: [28119636](https://pubmed.ncbi.nlm.nih.gov/28119636/)]
24. Knox M, Lentini J, Cummings T, McGrady A, Whearty K, Sancrant L. Game-based biofeedback for paediatric anxiety and depression. *Ment Health Fam Med* 2011 Sep;8(3):195-203 [FREE Full text] [Medline: [22942901](https://pubmed.ncbi.nlm.nih.gov/22942901/)]
25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009 Oct;62(10):e1-34 [FREE Full text] [doi: [10.1016/j.jclinepi.2009.06.006](https://doi.org/10.1016/j.jclinepi.2009.06.006)] [Medline: [19631507](https://pubmed.ncbi.nlm.nih.gov/19631507/)]
26. Burt J, Ravid EN, Bradford S, Fisher NJ, Zeng Y, Chomiak T, et al. The effects of music-contingent gait training on cognition and mood in parkinson disease: a feasibility study. *Neurorehabil Neural Repair* 2020 Jan;34(1):82-92. [doi: [10.1177/1545968319893303](https://doi.org/10.1177/1545968319893303)] [Medline: [31878824](https://pubmed.ncbi.nlm.nih.gov/31878824/)]
27. Venuturupalli RS, Chu T, Vicari M, Kumar A, Fortune N, Spielberg B. Virtual reality-based biofeedback and guided meditation in rheumatology: a pilot study. *ACR Open Rheumatol* 2019 Dec;1(10):667-675 [FREE Full text] [doi: [10.1002/acr2.11092](https://doi.org/10.1002/acr2.11092)] [Medline: [31872189](https://pubmed.ncbi.nlm.nih.gov/31872189/)]
28. Kotwas I, McGonigal A, Khalfa S, Bastien-Toniazzo M, Bartolomei F, Micoulaud-Franchi J. A case-control study of skin conductance biofeedback on seizure frequency and emotion regulation in drug-resistant temporal lobe epilepsy. *Int J Psychophysiol* 2018 Jan;123:103-110. [doi: [10.1016/j.ijpsycho.2017.10.005](https://doi.org/10.1016/j.ijpsycho.2017.10.005)] [Medline: [29042154](https://pubmed.ncbi.nlm.nih.gov/29042154/)]
29. Gaggioli A, Pallavicini F, Morganti L, Serino S, Scaratti C, Briguglio M, et al. Experiential virtual scenarios with real-time monitoring (interreality) for the management of psychological stress: a block randomized controlled trial. *J Med Internet Res* 2014 Jul 08;16(7):e167 [FREE Full text] [doi: [10.2196/jmir.3235](https://doi.org/10.2196/jmir.3235)] [Medline: [25004803](https://pubmed.ncbi.nlm.nih.gov/25004803/)]
30. Pallavicini F, Algeri D, Repetto C, Gorini A, Riva G. Biofeedback, virtual reality and mobile phones in the treatment of Generalized Anxiety Disorder (GAD): a phase-2 controlled clinical trial. *J Cyber Ther Rehabil* 2009;2(4):315-327 [FREE Full text]
31. Schuurmans AAT, Nijhof KS, Engels RCME, Granic I. Using a videogame intervention to reduce anxiety and externalizing problems among youths in residential care: an initial randomized controlled trial. *J Psychopathol Behav Assess* 2018;40(2):344-354 [FREE Full text] [doi: [10.1007/s10862-017-9638-2](https://doi.org/10.1007/s10862-017-9638-2)] [Medline: [29937625](https://pubmed.ncbi.nlm.nih.gov/29937625/)]
32. Li F, Xiong Y. Application of music therapy combined with computer biofeedback in the treatment of anxiety disorders. In: *Proceedings of the 8th International Conference on Information Technology in Medicine and Education*. Fuzhou, China; 2016 Presented at: 8th International Conference on Information Technology in Medicine and Education; December 23-25; Fuzhou, China p. 90-93. [doi: [10.1109/itme.2016.0029](https://doi.org/10.1109/itme.2016.0029)]
33. Vitasari P, Wahab M, Othman A, Awang M. The effectiveness of biofeedback training among university students. Malaysia; 2009 Presented at: IEEE Student Conference on Research and Development (SCORED); October 15-17; Serdang, Malaysia p. 551-554. [doi: [10.1109/scored.2009.5442939](https://doi.org/10.1109/scored.2009.5442939)]
34. Kuriakose S, Lahiri U. Design of a physiology-sensitive vr-based social communication platform for children with autism. *IEEE Trans Neural Syst Rehabil Eng* 2017 Aug;25(8):1180-1191. [doi: [10.1109/tnsre.2016.2613879](https://doi.org/10.1109/tnsre.2016.2613879)]
35. van Rooij M, Lobel A, Harris O, Smit N, Granic I. DEEP: a biofeedback virtual reality game for children at-risk for anxiety. San Jose, California, USA; 2016 Presented at: CHI Conference on Human Factors in Computing Systems; May 7-12; San Jose, California, USA. [doi: [10.1145/2851581.2892452](https://doi.org/10.1145/2851581.2892452)]
36. McKenna K, Gallagher KAS, Forbes PW, Ibeziako P. Ready, set, relax: biofeedback-assisted relaxation training (BART) in a pediatric psychiatry consultation service. *Psychosomatics* 2015;56(4):381-389. [doi: [10.1016/j.psym.2014.06.003](https://doi.org/10.1016/j.psym.2014.06.003)] [Medline: [25556570](https://pubmed.ncbi.nlm.nih.gov/25556570/)]
37. Scharff L, Marcus DA, Masek BJ. A controlled study of minimal-contact thermal biofeedback treatment in children with migraine. *J Pediatr Psychol* 2002 Mar;27(2):109-119. [doi: [10.1093/jpepsy/27.2.109](https://doi.org/10.1093/jpepsy/27.2.109)] [Medline: [11821495](https://pubmed.ncbi.nlm.nih.gov/11821495/)]
38. Krusemark EA, Li W. Enhanced olfactory sensory perception of threat in anxiety: an event-related fMRI study. *Chemosens Percept* 2012 Mar 01;5(1):37-45 [FREE Full text] [doi: [10.1007/s12078-011-9111-7](https://doi.org/10.1007/s12078-011-9111-7)] [Medline: [22866182](https://pubmed.ncbi.nlm.nih.gov/22866182/)]

39. Ihmig FR, Neurohr-Parakenings F, Schäfer SK, Lass-Hennemann J, Michael T. On-line anxiety level detection from biosignals: machine learning based on a randomized controlled trial with spider-fearful individuals. *PLoS One* 2020;15(6):e0231517 [FREE Full text] [doi: [10.1371/journal.pone.0231517](https://doi.org/10.1371/journal.pone.0231517)] [Medline: [32574167](https://pubmed.ncbi.nlm.nih.gov/32574167/)]
40. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. *Front Robot AI* 2016 Dec 19;3:74. [doi: [10.3389/frobt.2016.00074](https://doi.org/10.3389/frobt.2016.00074)]
41. Burdea G, Coiffet P. Virtual reality technology. *Presence (Camb)* 2003 Dec 01;12(6):663-664. [doi: [10.1162/105474603322955950](https://doi.org/10.1162/105474603322955950)]
42. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol Med* 2017 Oct;47(14):2393-2400 [FREE Full text] [doi: [10.1017/S003329171700040X](https://doi.org/10.1017/S003329171700040X)] [Medline: [28325167](https://pubmed.ncbi.nlm.nih.gov/28325167/)]
43. Gorini A, Riva G. Virtual reality in anxiety disorders: the past and the future. *Expert Rev Neurother* 2008 Feb 09;8(2):215-233. [doi: [10.1586/14737175.8.2.215](https://doi.org/10.1586/14737175.8.2.215)] [Medline: [18271709](https://pubmed.ncbi.nlm.nih.gov/18271709/)]
44. Powers MB, Emmelkamp PMG. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. *J Anxiety Disord* 2008;22(3):561-569. [doi: [10.1016/j.janxdis.2007.04.006](https://doi.org/10.1016/j.janxdis.2007.04.006)] [Medline: [17544252](https://pubmed.ncbi.nlm.nih.gov/17544252/)]
45. Botella C, García-Palacios A, Villa H, Baños RM, Quero S, Alcañiz M, et al. Virtual reality exposure in the treatment of panic disorder and agoraphobia: a controlled study. *Clin Psychol Psychother* 2007 May;14(3):164-175. [doi: [10.1002/cpp.524](https://doi.org/10.1002/cpp.524)]
46. Rothbaum BO, Hodges LF, Ready D, Graap K, Alarcon RD. Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *J Clin Psychiatry* 2001 Aug 15;62(8):617-622. [doi: [10.4088/JCP.v62n0808](https://doi.org/10.4088/JCP.v62n0808)]
47. Rodriguez N. Identifying accessibility conditions for children with multiple disabilities: a virtual reality wheelchair simulator. 2018 Presented at: IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); October 16-18; Munich, Germany. [doi: [10.1109/ismar-adjunct.2018.00107](https://doi.org/10.1109/ismar-adjunct.2018.00107)]
48. Zyda M. From visual simulation to virtual reality to games. *Computer* 2005 Sep;38(9):25-32. [doi: [10.1109/mc.2005.297](https://doi.org/10.1109/mc.2005.297)]
49. Schild J, Misztal S, Roth B, Flock L, Luiz T, Lerner D, et al. Applying multiuser virtual reality to collaborative medical training. In: *Proceeding of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Reutlingen; 2018 Presented at: IEEE Conference on Virtual Reality and 3D User Interfaces; March 18-22; Tuebingen/Reutlingen, Germany p. 775-776. [doi: [10.1109/vr.2018.8446160](https://doi.org/10.1109/vr.2018.8446160)]
50. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. *Cochrane Database Syst Rev* 2017 Nov 20;11:CD008349 [FREE Full text] [doi: [10.1002/14651858.CD008349.pub4](https://doi.org/10.1002/14651858.CD008349.pub4)] [Medline: [29156493](https://pubmed.ncbi.nlm.nih.gov/29156493/)]
51. Michael D, Chen S. *Serious Games: Games That Educate, Train and Inform*. Boston, Massachusetts: Thomson Course Technology; 2006.
52. Stapleton A. *Serious games: serious opportunities*. 2004 Presented at: Australian Game Developers Conference, Academic Summit; December 2; Melbourne, Australia p. 2.
53. Read JL, Shortell SM. Interactive games to promote behavior change in prevention and treatment. *JAMA* 2011 Apr 27;305(16):1704-1705. [doi: [10.1001/jama.2011.408](https://doi.org/10.1001/jama.2011.408)] [Medline: [21447802](https://pubmed.ncbi.nlm.nih.gov/21447802/)]
54. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. *Comput Educ* 2012 Sep;59(2):661-686. [doi: [10.1016/j.compedu.2012.03.004](https://doi.org/10.1016/j.compedu.2012.03.004)]
55. Andrews A. *Serious games for psychological health education*. 2011 Presented at: International Conference on Virtual and Mixed Reality; July 9-14; Orlando, Florida p. 3-10. [doi: [10.1007/978-3-642-22024-1\\_1](https://doi.org/10.1007/978-3-642-22024-1_1)]
56. Dondlinger M. Educational video game design: a review of the literature. *J Appl Educ Technol* 2007;4(1):21-31 [FREE Full text]
57. Leahy A, Clayman C, Mason I, Lloyd G, Epstein O. Computerised biofeedback games: a new method for teaching stress management and its use in irritable bowel syndrome. *J R Coll Physicians Lond* 1998;32(6):552-556. [Medline: [9881312](https://pubmed.ncbi.nlm.nih.gov/9881312/)]
58. Hung J, Yu M, Chang K, Lee H, Hsieh Y, Chen P. Feasibility of using Tetrax biofeedback video games for balance training in patients with chronic hemiplegic stroke. *PM R* 2016 Oct;8(10):962-970. [doi: [10.1016/j.pmrj.2016.02.009](https://doi.org/10.1016/j.pmrj.2016.02.009)] [Medline: [26968609](https://pubmed.ncbi.nlm.nih.gov/26968609/)]
59. Gough C. Distribution of video gamers worldwide in 2017, by age group and gender. Statista. 2017. URL: <https://www.statista.com/statistics/722259/world-gamers-by-age-and-gender/> [accessed 2021-01-04]
60. Gough C. Share of active mobile gamers worldwide in 2020, by age group. Statista. 2020. URL: <https://www.statista.com/statistics/748089/number-mobile-gamers-world-platform/> [accessed 2021-01-04]
61. Payne R, Donaghy M. *Payne's Handbook of Relaxation Techniques 4th Edition*. Amsterdam: Elsevier Health Sciences; 2010.
62. Eppley KR, Abrams AI, Shear J. Differential effects of relaxation techniques on trait anxiety: a meta-analysis. *J Clin Psychol* 1989 Nov;45(6):957-974. [doi: [10.1002/1097-4679\(198911\)45:6<957::aid-jclp2270450622>3.0.co;2-q](https://doi.org/10.1002/1097-4679(198911)45:6<957::aid-jclp2270450622>3.0.co;2-q)]
63. Vavra KL, Janjic-Watrich V, Loerke K, Phillips LM, Norris SP, Macnab J. Visualization in science education. *Alta Sci Educ J* 2011;41(1):22-30.
64. Enzenhofer M, Bludau H, Komm N, Wild B, Mueller K, Herzog W, et al. Improvement of the educational process by computer-based visualization of procedures: randomized controlled trial. *J Med Internet Res* 2004 Jun 02;6(2):e16 [FREE Full text] [doi: [10.2196/jmir.6.2.e16](https://doi.org/10.2196/jmir.6.2.e16)] [Medline: [15249265](https://pubmed.ncbi.nlm.nih.gov/15249265/)]

65. Rusy LM, Weisman SJ. Complementary therapies for acute pediatric pain management. *Pediatr Clin North Am* 2000 Jun;47(3):589-599. [doi: [10.1016/s0031-3955\(05\)70227-3](https://doi.org/10.1016/s0031-3955(05)70227-3)]
66. Smith CA, Levett KM, Collins CT, Armour M, Dahlen HG, Sukanuma M. Relaxation techniques for pain management in labour. *Cochrane Database Syst Rev* 2018 Mar 28;3:CD009514 [FREE Full text] [doi: [10.1002/14651858.CD009514.pub2](https://doi.org/10.1002/14651858.CD009514.pub2)] [Medline: [29589650](https://pubmed.ncbi.nlm.nih.gov/29589650/)]
67. Pyati S, Gan TJ. Perioperative pain management. *CNS Drugs* 2007;21(3):185-211. [doi: [10.2165/00023210-200721030-00002](https://doi.org/10.2165/00023210-200721030-00002)] [Medline: [17338592](https://pubmed.ncbi.nlm.nih.gov/17338592/)]
68. Carroll D, Seers K. Relaxation for the relief of chronic pain: a systematic review. *J Adv Nurs* 1998 Mar 25;27(3):476-487. [doi: [10.1046/j.1365-2648.1998.00551.x](https://doi.org/10.1046/j.1365-2648.1998.00551.x)] [Medline: [9543032](https://pubmed.ncbi.nlm.nih.gov/9543032/)]
69. Crouch R, Alers V. *Occupational Therapy in Psychiatry and Mental Health*. West Sussex, UK: Wiley Blackwell; 2014:978-971.
70. Robb SL. Music assisted progressive muscle relaxation, progressive muscle relaxation, music listening, and silence: a comparison of relaxation techniques. *J Music Ther* 2000 Mar 01;37(1):2-21. [doi: [10.1093/jmt/37.1.2](https://doi.org/10.1093/jmt/37.1.2)] [Medline: [10806470](https://pubmed.ncbi.nlm.nih.gov/10806470/)]
71. Stark DP, House A. Anxiety in cancer patients. *Br J Cancer* 2000 Nov;83(10):1261-1267 [FREE Full text] [doi: [10.1054/bjoc.2000.1405](https://doi.org/10.1054/bjoc.2000.1405)] [Medline: [11044347](https://pubmed.ncbi.nlm.nih.gov/11044347/)]
72. Yohannes AM, Alexopoulos GS. Depression and anxiety in patients with COPD. *Eur Respir Rev* 2014 Sep 31;23(133):345-349 [FREE Full text] [doi: [10.1183/09059180.00007813](https://doi.org/10.1183/09059180.00007813)] [Medline: [25176970](https://pubmed.ncbi.nlm.nih.gov/25176970/)]
73. Quirk ME, Letendre AJ, Ciottone RA, Lingley JF. Anxiety in patients undergoing MR imaging. *Radiology* 1989 Feb;170(2):6-463. [doi: [10.1148/radiology.170.2.2911670](https://doi.org/10.1148/radiology.170.2.2911670)] [Medline: [2911670](https://pubmed.ncbi.nlm.nih.gov/2911670/)]

## Abbreviations

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-analysis

**VR:** virtual reality

*Edited by L Buis; submitted 25.11.20; peer-reviewed by S Chen, HJ Jeon, R Liu; comments to author 29.01.21; revised version received 16.02.21; accepted 31.03.21; published 22.04.21.*

*Please cite as:*

*Alneyadi M, Drissi N, Almeqbaali M, Ouhbi S*

*Biofeedback-Based Connected Mental Health Interventions for Anxiety: Systematic Literature Review*

*JMIR Mhealth Uhealth* 2021;9(4):e26038

URL: <https://mhealth.jmir.org/2021/4/e26038>

doi: [10.2196/26038](https://doi.org/10.2196/26038)

PMID: [33792548](https://pubmed.ncbi.nlm.nih.gov/33792548/)

©Mahra Alneyadi, Nidal Drissi, Mariam Almeqbaali, Sofia Ouhbi. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 22.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

# Impact of Mobile Health Devices for the Detection of Atrial Fibrillation: Systematic Review

Tom E Biersteker<sup>1</sup>, MD; Martin J Schalijs<sup>1</sup>, MD, PhD; Roderick W Treskes<sup>1</sup>, MD, PhD

Leiden University Medical Center, Leiden, Netherlands

**Corresponding Author:**

Martin J Schalijs, MD, PhD

Leiden University Medical Center

Albinusdreef 2

Leiden

Netherlands

Phone: 31 715262020

Email: [m.j.schalijs@lumc.nl](mailto:m.j.schalijs@lumc.nl)

## Abstract

**Background:** Atrial fibrillation (AF) is the most common arrhythmia, and its prevalence is increasing. Early diagnosis is important to reduce the risk of stroke. Mobile health (mHealth) devices, such as single-lead electrocardiogram (ECG) devices, have been introduced to the worldwide consumer market over the past decade. Recent studies have assessed the usability of these devices for detection of AF, but it remains unclear if the use of mHealth devices leads to a higher AF detection rate.

**Objective:** The goal of the research was to conduct a systematic review of the diagnostic detection rate of AF by mHealth devices compared with traditional outpatient follow-up. Study participants were aged 16 years or older and had an increased risk for an arrhythmia and an indication for ECG follow-up—for instance, after catheter ablation or presentation to the emergency department with palpitations or (near) syncope. The intervention was the use of an mHealth device, defined as a novel device for the diagnosis of rhythm disturbances, either a handheld electronic device or a patch-like device worn on the patient's chest. Control was standard (traditional) outpatient care, defined as follow-up via general practitioner or regular outpatient clinic visits with a standard 12-lead ECG or Holter monitoring. The main outcome measures were the odds ratio (OR) of AF detection rates.

**Methods:** Two reviewers screened the search results, extracted data, and performed a risk of bias assessment. A heterogeneity analysis was performed, forest plot made to summarize the results of the individual studies, and albatross plot made to allow the *P* values to be interpreted in the context of the study sample size.

**Results:** A total of 3384 articles were identified after a database search, and 14 studies with a 4617 study participants were selected. All studies but one showed a higher AF detection rate in the mHealth group compared with the control group (OR 1.00-35.71), with all RCTs showing statistically significant increases of AF detection (OR 1.54-19.16). Statistical heterogeneity between studies was considerable, with a *Q* of 34.1 and an  $I^2$  of 61.9, and therefore it was decided to not pool the results into a meta-analysis.

**Conclusions:** Although the results of 13 of 14 studies support the effectiveness of mHealth interventions compared with standard care, study results could not be pooled due to considerable clinical and statistical heterogeneity. However, smartphone-connectable ECG devices provide patients with the ability to document a rhythm disturbance more easily than with standard care, which may increase empowerment and engagement with regard to their illness. Clinicians must beware of overdiagnosis of AF, as it is not yet clear when an mHealth-detected episode of AF must be deemed significant.

(*JMIR Mhealth Uhealth* 2021;9(4):e26161) doi:[10.2196/26161](https://doi.org/10.2196/26161)

**KEYWORDS**

eHealth; mHealth; telemedicine; cardiology; atrial fibrillation; systematic review

## Introduction

Atrial fibrillation (AF) is the most commonly diagnosed arrhythmia [1]. It may be paroxysmal (present for 30 seconds

to 7 days), persistent (present for more than 7 days), or permanent [2]. Risk factors for AF are diverse and include advanced age, male gender, diabetes mellitus, hypertension, obesity, valvular disease, obstructive sleep apnea, heart failure, and previous myocardial infarction [3]. Among other symptoms,



AF can cause palpitations, dyspnea, and tiredness. Patients can, however, be asymptomatic [4].

The worldwide prevalence of AF is increasing. This increase has been attributed to an aging population and increased prevalence of cardiovascular risk factors [5]. A European study has shown that the number of patients with diagnosed AF is expected to increase from a prevalence of 2.3% in 2010 to 3.5% to 4.3% in 2050 [6]. Due to an increased risk of stroke, AF is associated with increased risk of mortality [7]. Compared with patients with sinus rhythm, those with AF are found to have a 2.4-fold risk of stroke, and the risk of ischemic heart disease and development of chronic kidney disease are both increased 1.6-fold [8].

Early diagnosis of AF and prophylactic treatment for ischemic stroke with oral anticoagulants is therefore important, whether the AF is paroxysmal, persistent, or permanent and symptomatic or silent [2]. Moreover, it has been demonstrated that excessive supraventricular ectopic activity, defined as the presence of either  $\geq 30$  premature atrial contractions (PACs) per hour daily or any runs of  $\geq 20$  PACs, increases the risk of stroke in patients with a CHA<sub>2</sub>DS<sub>2</sub>-VASc (congestive heart failure, hypertension, age  $\geq 75$  years, diabetes mellitus, stroke or transient ischemic attack [TIA], vascular disease, age 65 to 74 years, sex category) score of  $\geq 2$  by 2.4% [9].

Traditionally, patients are diagnosed with AF using a 12-lead electrocardiogram (ECG). In case of suspected paroxysmal AF, it is possible to perform prolonged monitoring via Holter registration. However, as paroxysmal AF is often silent and patients can have vast periods of sinus rhythm, diagnosing paroxysmal AF is a challenge [10].

Over the last decade, consumer grade health monitoring devices have been developed and marketed as beneficial for personal health monitoring [11]. Among those devices are several different smartphone connectable ECG devices. The majority are lead-I ECG devices, handheld instruments that register lead I of the ECG, measuring the electric current generated by the myocardium by using the fingers of the right and the left hand [12]. These devices are typically used for spot-checks. Another group of devices is meant for continuous monitoring and involve patches that stick to the chest and allow monitoring of the heart rate and rhythm continuously for up to 2 weeks [13]. Both groups of devices can be seen as mobile health (mHealth) devices and used for AF screening [12].

Studies have been done to assess the accuracy of mHealth devices compared with 12-lead ECGs. A recent systematic review suggests several mHealth devices are suitable in the use of detecting AF, based on the sensitivity and specificity of these devices [14]. However, it is still unclear if and to what extent the use of mHealth devices leads to higher detection rates of AF. Therefore, the objective of this systematic review is to evaluate studies comparing the detection rate of AF by mHealth devices with more traditional outpatient follow-up.

## Methods

### Literature Review and Definitions

A systematic literature review was conducted to evaluate the efficacy of mHealth devices using standard (traditional) care as the reference standard in people with an indication for follow-up for a suspected arrhythmia (eg, after catheter ablation or electrical cardioversion) or in cases of an acute emergency department presentation with (near) syncope or palpitations where no arrhythmia could be found at the time of presentation. The efficacy of mHealth was defined as the detection rate of AF by a smartphone-connectable ECG device, either a handheld electronic device or patch-like device attached to the study subject's chest or by requiring subject to send an ECG transtelephonically. Standard care was defined as follow-up via a general practitioner or regular outpatient clinic visit with a standard 12-lead ECG or Holter monitoring. This systematic review was conducted and reported by following the Cochrane Handbook for Systematic Reviews of Interventions [15].

### Eligibility Criteria

The eligibility criteria for studies to be included in this systematic review were as follows:

- Published studies comparing mHealth devices with standard care in patients with an indication for follow-up via ECG or Holter monitoring
- Studies with AF detection as a primary or secondary outcome measure
- Studies conducted in people aged 16 years and older reporting demographic data such as patient characteristics, study setting, sample size, and data points
- Studies performed in a clinical or outpatient setting
- Studies in patients without an internal cardioverter defibrillator, pacemaker, or ventricular assist device

Studies had to be published in English or Dutch to be selected. If a study has been indexed in multiple databases, only the PubMed version was included.

### Literature Search Strategy

The search strategy is presented in [Multimedia Appendix 1](#). No study design filters were applied, and all electronic databases were searched for articles from Jan 1, 2005, until February 19, 2020. The following databases were searched: Medline, Embase, PubMed, Web of Science, Emcare, Academic Search Premier, and the Cochrane Library. The search results were managed using EndNote X9 software (Clarivate Analytics). Relevant studies and reviews were manually searched to identify other possible relevant studies.

### Article Selection and Data Synthesis

A 2-stage process was used for inclusion in the review. Two reviewers (TB, RT) first independently screened all titles and abstracts of the identified studies to find potentially relevant studies. The same reviewers then assessed the full-text articles independently for the eligibility criteria. Any disagreements were resolved by consensus.

### Risk of Bias Assessment

Risk of bias was assessed with the RoB 2 (Risk of Bias 2) tool for randomized controlled trials (RCTs) and the ROBINS-I (Risk of Bias in Nonrandomized Studies of Interventions) tool for nonrandomized studies [16,17]. This is in accordance with the Cochrane Handbook's recommendations [15]. The risk of bias had 3 levels: low risk of bias, some concerns, and high risk of bias.

### Summary Measures

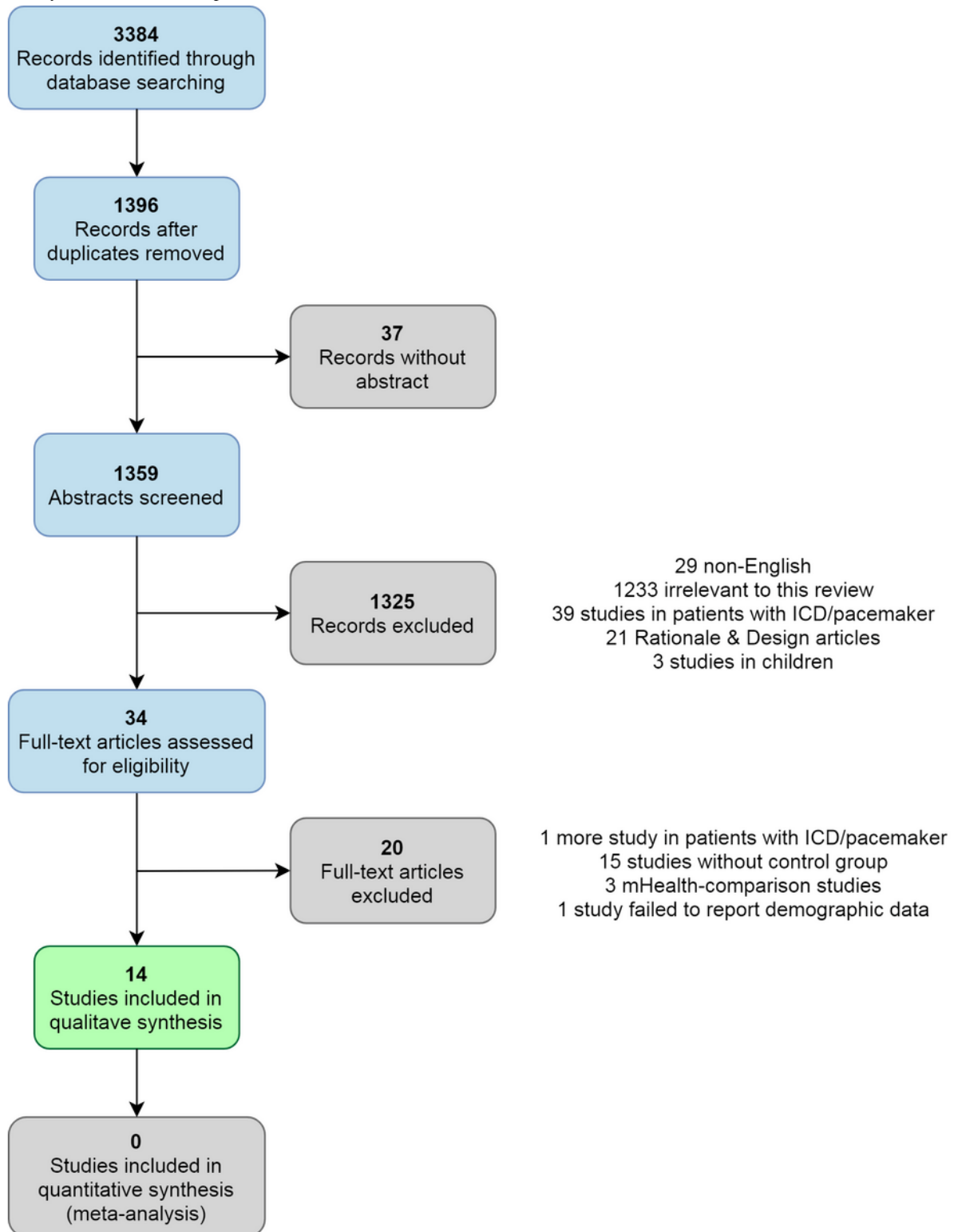
The primary outcome measure of this systematic review was the odds ratio (OR) of AF detection, comparing mHealth devices to standard care. The PATCH-ED (Patch Monitor in Patients With Unexplained Syncope After Initial Evaluation in the Emergency Department) and IPED (Investigation of Palpitations in the Emergency Department) study groups reported no events in the control groups [18,19]. Therefore, the Haldane correction was used [20]. A heterogeneity analysis between studies was performed with a chi-square test [15]. A forest plot was made to summarize the results of individual studies. Finally, an

albatross plot was made to allow the  $P$  values to be interpreted in the context of study sample size. The contour lines of albatross plots are formed by hypothetical effect sizes [21]. In this case, this concerns odds ratios due to the outcome being dichotomous. The forest and albatross plots were made in Matlab (The Mathworks Inc).

## Results

### Study Selection

As of October 19, 2020, a total of 3384 articles were obtained from the database searches. Two investigators (TB and RT) excluded 3350 studies based on the title and abstract. A total of 34 abstracts meeting the eligibility criteria were identified. After reviewing the full text, the reviewers chose 14 studies with a total of 4617 study subjects. The selection process is shown in Figure 1. The kappa statistic for interrater reliability was .81, showing substantial agreement between the 2 investigators [22].

**Figure 1.** Study search and selection process.

### Study Characteristics

The 14 selected studies consist of 8 cohort studies, 4 RCTs, and 2 case-control studies [18,19,23-34]. Table 1 shows participant and study characteristics. Study populations were heterogenous:

some studies included only patients without any history of AF, others included only patients with earlier documented AF. Participant genders varied between the study populations: 42% to 87% were male. Mean age varied from 44 to 73 years.

**Table 1.** Study characteristics.

Author, year, country	Study type	Patient characteristics	Sample size; drop out; mean age; male	Intervention	Control	Follow-up	Primary outcome
Liu et al (2010), China [23]	Prospective cross-sectional	Catheter ablation patients	92; 0 (0%); 54 y <sup>a</sup> ; 78% male	Transtelephonic ECG <sup>b</sup> once daily	24 h Holter+ at complaints	90 d <sup>c</sup>	AF <sup>d</sup> detection
Rosenberg et al (2013), US [24]	Prospective cross-sectional	Patients who are managed for AF, no definition was given	74; 0 (0%); 65 y; 55% male	ZioPatch	24 h Holter	14 d	AF detection
Barrett et al (2013), US [25]	Prospective cross-sectional	Outpatients with indication for Holter monitoring	146; 4 (2.7%); n/a <sup>e</sup> ; n/a	ZioPatch	24 h Holter	14 d	Arrhythmia detection
Hendrikx et al (2014), Sweden [26]	Prospective cross-sectional	Patients with unexplained palpitations or presyncope	95; 0 (0%); 54 y; 44% male	Zenikor twice daily + 24 h <sup>f</sup> Holter	24 h Holter	28 d	Arrhythmia detection
Kimura et al (2016), Japan [27]	Prospective cross-sectional	Catheter ablation patients	28; 2 (6.7%); 59 y; 87% male	CardioPhone twice daily	Monthly 24 h Holter	6 mo <sup>g</sup>	AF detection
Busch et al (2017), Germany [28]	Retrospective cross-sectional	Volunteers to join in an mHealth <sup>h</sup> study	1678; n/a; 51 y; 48% male	SensorMobile twice daily	Single 12-lead ECG	28 d	AF detection
Halcox et al (2017), UK [29]	Single center, open label RCT <sup>i</sup>	≥65 y patients without AF at a GP <sup>j</sup> practice	1001; 5 (0.5%); 73 y; 47% male	AliveCor Kardia twice a week	Follow-up at the GP	1 y	Time to diagnosis of AF
Hickey et al (2017), US [30]	Prospective, matched cohort study	Patients with a history of AF	46; 0 (0%); 55 y; 65% male	AliveCor Kardia once daily	Standard care (no added care)	6 mo	Atrial arrhythmia detection
Narasimha et al (2018), US [31]	Prospective cross-sectional	Patients with unexplained palpitations who underwent previous Holter monitoring	33; 5 (13.2%); 48 y; 42% male	AliveCor Kardia at complaints	External loop recorder	30 d	Arrhythmia detection
Reed et al (2018), Scotland [18]	Prospective, unmatched case-control study	≥16 y ER patients with unexplained syncope	689; 0 (0%); 67 y; 47% male	ZioPatch	Standard care (no added care)	14 d	Symptomatic rhythm detection
Reed et al (2019), Scotland [19]	Multicenter, open label RCT	≥16 y ER patients with unexplained palpitations or (pre)syncope	240; 2 (0.8%); 40 y; 44% male	Alivecor Kardia at complaints	Standard care (no added care)	90 d	Symptomatic rhythm detection
Goldenthal et al (2019), US [32]	Single center, open label RCT	Patients with documented AF, undergoing ablation or ECV <sup>k</sup>	238; 5 (2.1%); 61 y; 76% male	AliveCor Kardia daily and at complaints	Standard care (no added care)	6 mo	AF detection
Karunadas et al (2019), India [33]	Prospective cross-sectional	Admitted patients to cardiology ward who required monitoring	141; 0 (0%); 44 y; 53% male	WebCardio (patch)	24 h Holter	1 d	Arrhythmia detection
Kaura et al (2019), UK [34]	Multicenter, open label RCT	Non-AF patients with non-lacunar stroke or TIA <sup>l</sup>	116; 26 (22.4%); 70 y; 47% male	ZioPatch	24 h Holter	14 d	AF detection

<sup>a</sup>y: year.<sup>b</sup>ECG: electrocardiogram.<sup>c</sup>d: day.<sup>d</sup>AF: atrial fibrillation.<sup>e</sup>Not applicable.<sup>f</sup>h: hour.<sup>g</sup>mo: month.<sup>h</sup>mHealth: mobile health.<sup>i</sup>RCT: randomized controlled trial.<sup>j</sup>GP: general practice.

<sup>k</sup>ECV: electrical cardioversion.

<sup>l</sup>TIA: transient ischemic attack.

A total of 9 studies used handheld devices such as the Kardia (AliveCor Inc) or Zenicor-ECG (Zenicor Medical Systems AB) as an intervention, while 5 studies used a patch such as the Zio (iRhythm Technologies Inc), which was placed on the participant's chest [13,35,36]. The duration of the intervention was 1 to 14 days for studies with patches and 28 days to 1 year for studies with handheld devices. All studies published data about AF detection, although AF detection was the primary outcome in only 6 studies. A total of 4 studies used detection of any arrhythmia (AF, atrial flutter, supraventricular or ventricular tachycardia, sinus pauses of more than 3 seconds, and second- and third-degree atrioventricular blocks), and 2 other studies reported symptomatic arrhythmias as the primary outcome; 1 study used atrial arrhythmia detection and the final study reported the time to AF diagnosis as the primary outcome. One study reported a composite endpoint of AF, ventricular tachycardia, and sinus pauses of more than 3 seconds instead [25].

A total of 6 studies used 24-hour Holter monitoring as standard care, with 1 study adding another 24-hour Holter monitoring when study patients experienced an episode of palpitations and another study adding another 24-hour Holter monitoring every month, 6 times in total. However, 5 studies only saw patients back in the outpatient clinic or general practitioner. One study used an external loop recorder as standard care, activated at complaints during the entire follow-up duration, and the final study documented one extra standard ECG as standard care. Holter timing was at the start of the study in 4 of 6 studies that used Holter monitoring. In the other 2 studies, the timing of the Holter monitoring was unclear.

## Study Results

Table 2 shows the number of events throughout the studies. The individual study results are shown in a forest plot (Figure 2) but not pooled due to the considerable clinical and statistical heterogeneity. To show the *P* values in the context of the study sample size, an albatross plot is presented (Figure 3).

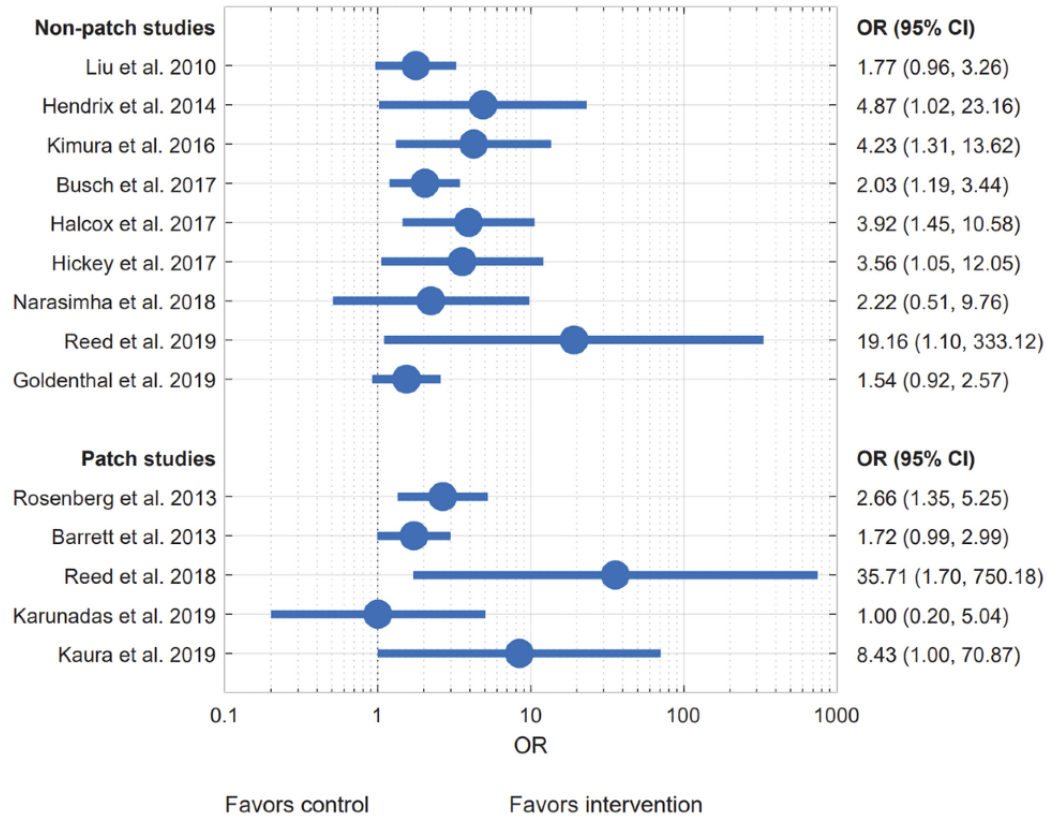
**Table 2.** Study outcomes.

Author	Sample size, n	Intervention group, n	Control group, n	Events (intervention), n (%)	Events (control), n (%)	Odds ratio (95% CI)
<b>Nonpatch studies</b>						
Liu et al, 2010 [23]	92	— <sup>a</sup>	—	39 (42.4)	27 (29.2)	1.77 (0.96-3.26)
Hendriks et al, 2014 [26]	95	—	—	9 (9.5)	2 (2.1)	4.87 (1.02-23.16)
Kimura et al, 2016 [27]	28	—	—	15 (53.6)	6 (21.4)	4.23 (1.31-13.62)
Busch et al, 2017 [28]	1678	—	—	42 (2.6)	21 (1.3)	2.03 (1.19-3.44)
Halcox et al, 2017 [29]	1001	500	501	19 (3.8)	5 (1.0)	3.92 (1.45-10.58)
Hickey et al, 2017 [30]	46	23	23	14 (60.9)	7 (30.4)	3.56 (1.05-12.05)
Narasimha et al, 2018 [31]	33	—	—	6 (18.2)	3 (9.1)	2.22 (0.51-9.76)
Reed et al, 2019 [19]	240	124	116	9 (7.3)	0 (0)	19.16 <sup>b</sup> (1.10-333.12)
Goldenthal et al, 2019 [32]	238	115	123	58 (50.4)	49 (41.5)	1.54 (0.92-2.57)
<b>Patch studies</b>						
Rosenberg et al, 2013 [24]	74	—	—	38 (51.3)	21 (28.4)	2.66 (1.35-5.26)
Barrett et al, 2013 [25]	146	—	—	41 (28.1)	27 (18.5)	1.72 (0.99-2.99)
Reed et al, 2018 [18]	689	86	603	2 (2.3)	0 (0)	35.71 <sup>b</sup> (1.70-750.18)
Karunadas et al, 2019 [33]	141	—	—	3 (2.1)	3 (2.1)	1.00 (0.20-5.04)
Kaura et al, 2019 [34]	116	56	60	7 (16.3)	1 (2.1)	8.43 (1.00-70.87)

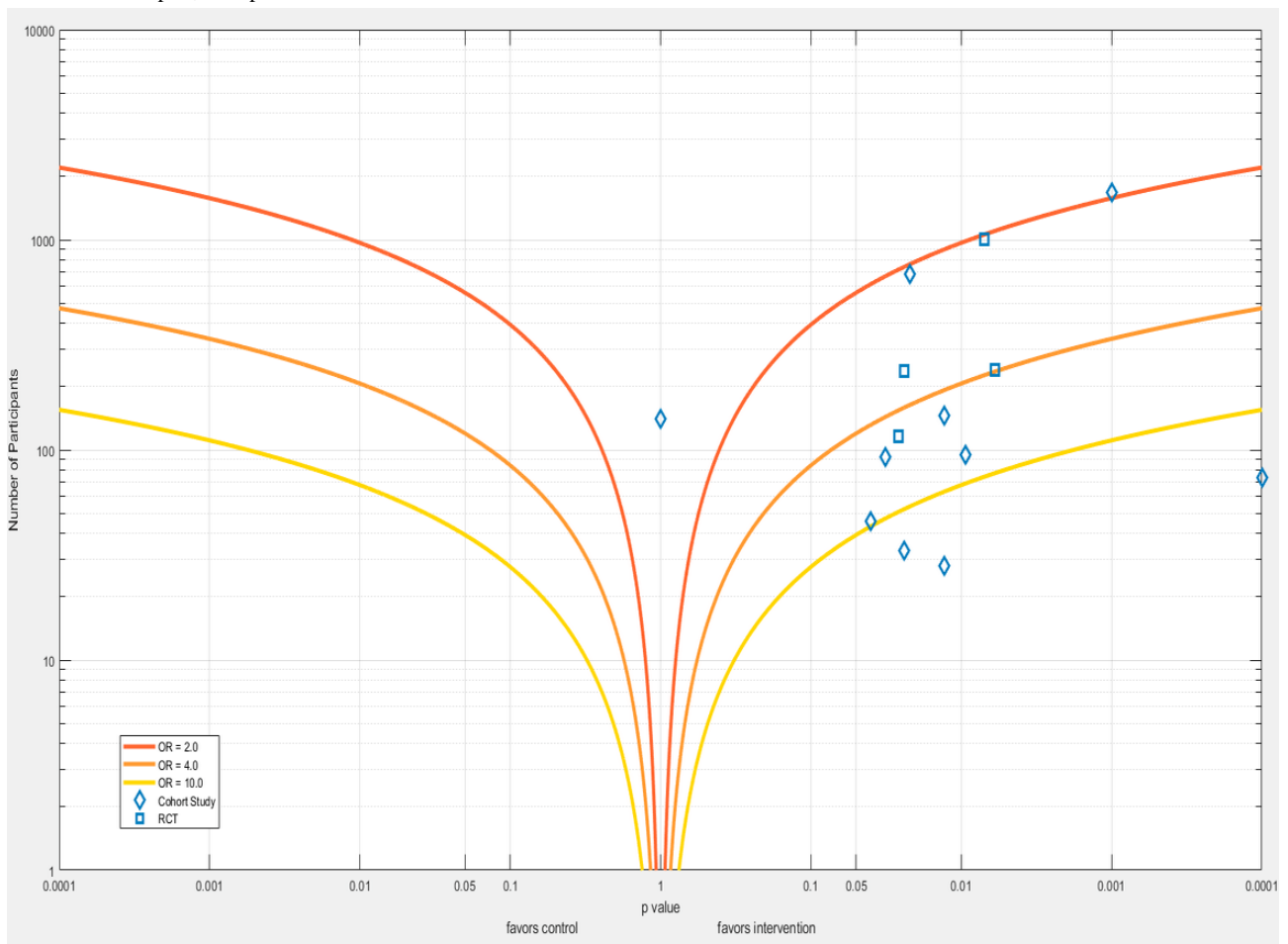
<sup>a</sup>Not applicable.

<sup>b</sup>Haldane correction applied.

**Figure 2.** Forest plot of the study results. No pooling due to heterogeneity.



**Figure 3.** Albatross plot, with plotted odds ratio lines.



All studies showed a higher AF detection rate in the mHealth group compared with the control group except the study by Karunadas, which showed an equal number of events (3; 2.1%) in both groups [33]. This study used an mHealth patch for 1 day and compared it to Holter monitoring performed on the same day. The 24-hour to 72-hour patch data have been disregarded for the analysis.

All RCTs showed a statistically significant improvement of AF detection with mHealth devices. ORs were 3.92 (95% CI 1.45-10.58) for the REHEARSE-AF (Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation) trial, 19.16 (95% CI 1.10-333.12) for IPED, 1.54 (95% CI 0.92-2.57) in the iHeart (Information Technology Approach to Implementing Depression Treatment in Cardiac Patients) trial, and 8.43 (95% CI 1.00-70.87) in the EPACS (Early Prolonged Ambulatory Cardiac Monitoring in Stroke) trial.

### Statistical Heterogeneity

The 14 selected studies showed a variety of populations, interventions, and outcomes and are therefore considerably clinically heterogenic. A chi-square test was conducted to assess statistical heterogeneity, which showed a Q of 34.1 and an  $I^2$  of 61.9, and therefore the studies show considerable statistical heterogeneity.

### Quality Appraisal

Figure 4 presents the generic risk of bias, assessed with the RoB 2 and ROBINS-I tools. In the selected RCTs, blinding of participants was not possible due to the nature of the intervention. Of all selected RCTs, one had a high risk of bias on the outcome data. Kaura et al [34] reported a dropout of 22.4% and did not address this data in the report. This was also true for the RCT by Goldenthal et al [32], but the dropout in this trial was just 2.1%. As for allocation concealment in the trial carried out by Halcox et al [29], no clarity was provided in the method section of the paper.

**Figure 4.** Risk of bias assessment. Randomized trials were assessed with the ROB 2 (Risk of Bias 2) tool, while ROBINS-I was used for nonrandomized studies. ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions.



Of the nonrandomized studies, the studies by Liu et al [23], Rosenberg et al [24], Hendrikx et al [26], Kimura et al [27], and Hickey et al [30] were scored as strong. Several studies showed an intermediate risk of bias. Barrett et al [18] reported no baseline characteristics, and Holter timing was unclear. Narasimha et al [25] reported a dropout of 13.2% but performed separate per-protocol and intention-to-treat analyses. Reed et al [31] used unmatched cohorts with several parameters not being known or stated. Also, there was a time interval of 7 to 8 years between gathering of the data in the intervention and control cohorts.

Two studies showed a high risk of bias. Busch et al [28] used data from a registry, in which the study subjects were volunteers willing to participate in an mHealth study. Karunadas et al [33] reported no baseline characteristics, and only WebCardio data from the first 24 hours were used. The 24-hour to 72-hour data, although gathered, were not reported.

## Discussion

### Summary of Evidence

The main finding of this systematic review of 14 studies is the increased AF detection rate when using mHealth devices compared with standard follow-up. Moreover, the 4 RCTs included all showed a statistically significant difference. However, there was a considerable clinical and statistical interstudy heterogeneity. The results of all studies but one show that mHealth devices lead to an increased detection of AF.

An argument can be made that conducting more (spot) measurements will automatically lead to more diagnoses of any illness. However, as AF is often only present for a short period of time and untraceable once sinus rhythm is restored, the clinical implications of the opportunity for conducting more spot measurements could be of importance with regard to stroke risk, for example. Following standard care does not allow patients to record their ECG without a delay, as they must visit their care provider or call an ambulance. Meanwhile, a paroxysm of AF may already have disappeared. Smartphone-connectable ECG devices could therefore provide patients with the opportunity to act immediately by documenting their rhythm disturbance. This is not only true for AF but also for other paroxysmal arrhythmias.

Although both handheld devices and patches lead to an increased AF detection rate, there may be a different use case to both groups of devices. Patches could be seen as prolonged Holter monitoring. The Zio patch can remain on the body for up to 14 days [13]. Handheld devices are used to do spot measurements for a longer period of time and can therefore only be used for screening or in patients with complaints that could fit with a rhythm disturbance. Therefore, the benefit of patches over handheld devices is that asymptomatic rhythm disturbances may be diagnosed with the use of a patch, although patient-triggered recordings with handheld ECG devices may be a more viable solution when a longer period of follow-up is indicated.

### Potential of mHealth for Population-Based Screening

Smartphone-connectable ECG devices cannot only be used in patients with a suspected paroxysmal rhythm disturbance but also for screening purposes. As stroke has been found to be the first symptom of AF in 37% of patients aged younger than 75 years with no history of cardiovascular diseases, secondary prevention in the form of screening risk groups for AF de novo may be of clinical relevance [37]. When it comes to screening for AF, there are several possibilities. Individuals can be screened regardless of medical history (systematic screening), on presenting to a physician for issues unrelated to AF (systematic opportunistic screening), or based on the presence of AF-associated risk factors (targeted screening). A recent meta-analysis has shown opportunistic screening, with a number needed to screen of 170, to be a likely cost-effective use of resources [38]. However, the number needed to screen varies between age groups and is found to be lowest, 83, in patients aged older than 65 years, against 926 for ages 60 to 64 years and 1089 for patients aged younger than 60 years, and therefore screening might be most opportune in people aged older than

65 years [39]. A very recent study using a Monte Carlo simulation to assess the cost-effectiveness of screening for AF with mHealth devices using 30,000 patients per CHA<sub>2</sub>DS<sub>2</sub>-VASc score (1-9) has found this type of screening to cause increased health care costs but a reduction in the incidence of stroke [40]. Several mHealth studies have used a systematic opportunistic screening approach such as screening for AF with handheld devices in individuals who visit pharmacies or those who visit their general practitioner for a flu vaccination [41-45]. These studies have all concluded handheld smartphone-connectable ECG devices to be viable screening tools.

### Clinical Implications

In this era of mHealth, patients are increasingly able to take (spot) measurements by using smartphone-connectable ECG devices, as those devices are commercially available. However, no consensus exists within the scientific community whether each episode of AF should be seen as clinically significant. AF is traditionally defined as an irregular arrhythmia without visible P waves lasting 30 seconds or more or documented on a standard 10-second 12-lead ECG [46]. The Kardia and other devices that register a lead-I ECG document a period of 30 seconds [35]. However, the clinical significance of a short paroxysm of AF is debated. Looking at AF ablation patients, it is known that the quality of life response is proportional to the burden rather than to a short-lived event and the AF burden is also a better predictor for stroke risk compared solely with a history of AF [47,48]. A recent study in patients with pacemakers tested various AF episode duration thresholds and found that patients with initial AF events up to 3.8 hours only had a median AF burden of 0.2% compared with 9.5% for those with initial AF episodes of more than 3.8 hours. This was a statistically significant difference with a *P* value of <.0001 [49].

### Limitations

Due to considerable clinical and statistic heterogeneity, with an *I*<sup>2</sup> of 61.9, the results of the included studies could not be pooled into a meta-analysis. The study populations varied from healthy adults to patients with an extensive history of AF, interventions ranged from short-term follow-up with a patch to long-term follow-up with a handheld device, and primary outcomes were also diverse. These differences led to a wide spread in the number of detected cases of AF, from 1% to 3% in the study by Busch et al [28] to 30% to 61% in the study by Hickey et al [30]. Instead of performing a meta-analysis, a forest plot without a diamond and an albatross plot were made. Furthermore, participants in RCTs could not be blinded due to the nature of the intervention. This is a small problem, however, since a diagnosis of AF is not a subjective end point.

### Conclusion

This systematic review reflects on 14 studies with different populations, interventions, and (primary) outcomes. A total of 13 studies found an increased number of AF diagnoses with the use of an mHealth intervention compared with standard care, with the remaining study by Karunadas et al [33] showing equal effectiveness. All 4 RCTs showed a statistically significant result in favor of the mHealth intervention. Due to considerable clinical and statistical heterogeneity, individual study results



could not be pooled into a meta-analysis, and as a result, it cannot be concluded that those mHealth interventions are effective in certain populations or every population. However, smartphone-connectable ECG devices provide patients with the ability to document a rhythm disturbance more easily than with standard care, and with the introduction of more mHealth devices and specifically devices that can diagnose AF like the Apple Watch (Apple Inc) and Move ECG (Withings) [50,51],

this is unlikely to change. With increased patient expectations and the increased empowerment and engagement with regard to their illness that mHealth devices may provide [52], future patients may request mHealth to be a part of their standard follow-up. However, as it is not yet clear when an mHealth-detected episode of AF should be deemed significant [48], clinicians must beware of overdiagnosis of AF and, sequentially, overtreatment with oral anticoagulants.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Search strategy.

[[DOCX File , 12 KB - mhealth\\_v9i4e26161\\_app1.docx](#) ]

## 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](https://doi.org/10.1161/CIRCULATIONAHA.113.005119)] [Medline: [24345399](https://pubmed.ncbi.nlm.nih.gov/24345399/)]
2. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 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 Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2014 Dec 02;64(21):e1-e76 [FREE Full text] [doi: [10.1016/j.jacc.2014.03.022](https://doi.org/10.1016/j.jacc.2014.03.022)] [Medline: [24685669](https://pubmed.ncbi.nlm.nih.gov/24685669/)]
3. Rienstra M, McManus DD, Benjamin EJ. Novel risk factors for atrial fibrillation: useful for risk prediction and clinical decision making? *Circulation* 2012 May 22;125(20):e941-e946 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.112.112920](https://doi.org/10.1161/CIRCULATIONAHA.112.112920)] [Medline: [22615425](https://pubmed.ncbi.nlm.nih.gov/22615425/)]
4. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016 Oct 07;37(38):2893-2962. [doi: [10.1093/eurheartj/ehw210](https://doi.org/10.1093/eurheartj/ehw210)] [Medline: [27567408](https://pubmed.ncbi.nlm.nih.gov/27567408/)]
5. Pistoia F, Sacco S, Tiseo C, Degan D, Ornello R, Carolei A. The epidemiology of atrial fibrillation and stroke. *Cardiol Clin* 2016 May;34(2):255-268. [doi: [10.1016/j.ccl.2015.12.002](https://doi.org/10.1016/j.ccl.2015.12.002)] [Medline: [27150174](https://pubmed.ncbi.nlm.nih.gov/27150174/)]
6. Stefansdottir H, Aspelund T, Gudnason V, Arnar DO. Trends in the incidence and prevalence of atrial fibrillation in Iceland and future projections. *Europace* 2011 Aug;13(8):1110-1117. [doi: [10.1093/europace/eur132](https://doi.org/10.1093/europace/eur132)] [Medline: [21551478](https://pubmed.ncbi.nlm.nih.gov/21551478/)]
7. Lip GYH, Tse HF, Lane DA. Atrial fibrillation. *Lancet* 2012 Feb 18;379(9816):648-661. [doi: [10.1016/S0140-6736\(11\)61514-6](https://doi.org/10.1016/S0140-6736(11)61514-6)] [Medline: [22166900](https://pubmed.ncbi.nlm.nih.gov/22166900/)]
8. Odotayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016 Sep 06;354:i4482 [FREE Full text] [doi: [10.1136/bmj.i4482](https://doi.org/10.1136/bmj.i4482)] [Medline: [27599725](https://pubmed.ncbi.nlm.nih.gov/27599725/)]
9. Larsen BS, Kumarathurai P, Falkenberg J, Nielsen OW, Sajadieh A. Excessive atrial ectopy and short atrial runs increase the risk of stroke beyond incident atrial fibrillation. *J Am Coll Cardiol* 2015 Jul 21;66(3):232-241 [FREE Full text] [doi: [10.1016/j.jacc.2015.05.018](https://doi.org/10.1016/j.jacc.2015.05.018)] [Medline: [26184616](https://pubmed.ncbi.nlm.nih.gov/26184616/)]
10. Thakkar S, Bagarhatta R. Detection of paroxysmal atrial fibrillation or flutter in patients with acute ischemic stroke or transient ischemic attack by Holter monitoring. *Indian Heart J* 2014;66(2):188-192 [FREE Full text] [doi: [10.1016/j.ihj.2014.02.009](https://doi.org/10.1016/j.ihj.2014.02.009)] [Medline: [24814113](https://pubmed.ncbi.nlm.nih.gov/24814113/)]
11. Treskes RW, Van Der Velde ET, Atsma DE, Schalij MJ. Redesigning healthcare: the 2.4 billion euro question—connecting smart technology to improve outcome of patients. *Neth Heart J* 2016 Jul;24(7-8):441-446. [doi: [10.1007/s12471-016-0834-6](https://doi.org/10.1007/s12471-016-0834-6)] [Medline: [27052894](https://pubmed.ncbi.nlm.nih.gov/27052894/)]
12. Duarte R, Stainthorpe A, Mahon J, Greenhalgh J, Richardson M, Nevitt S, et al. Lead-I ECG for detecting atrial fibrillation in patients attending primary care with an irregular pulse using single-time point testing: a systematic review and economic evaluation. *PLoS One* 2019;14(12):e0226671 [FREE Full text] [doi: [10.1371/journal.pone.0226671](https://doi.org/10.1371/journal.pone.0226671)] [Medline: [31869370](https://pubmed.ncbi.nlm.nih.gov/31869370/)]
13. iRhythmTech. URL: <http://www.irhythmtech.com> [accessed 2020-11-17]
14. Giebel GD, Gissel C. Accuracy of mHealth devices for atrial fibrillation screening: systematic review. *JMIR Mhealth Uhealth* 2019 Jun 16;7(6):e13641 [FREE Full text] [doi: [10.2196/13641](https://doi.org/10.2196/13641)] [Medline: [31199337](https://pubmed.ncbi.nlm.nih.gov/31199337/)]
15. Higgins J, Thomas J, Chandler J. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0. 2019. URL: <https://training.cochrane.org/handbook> [accessed 2020-11-17]

16. 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 28;366:l4898. [doi: [10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898)] [Medline: [31462531](https://pubmed.ncbi.nlm.nih.gov/31462531/)]
17. 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 12;355:i4919 [FREE Full text] [Medline: [27733354](https://pubmed.ncbi.nlm.nih.gov/27733354/)]
18. Reed MJ, Grubb NR, Lang CC, Gray AJ, Simpson K, MacRaid A, et al. Diagnostic yield of an ambulatory patch monitor in patients with unexplained syncope after initial evaluation in the emergency department: the PATCH-ED study. *Emerg Med J* 2018 Aug;35(8):477-485. [doi: [10.1136/emmermed-2018-207570](https://doi.org/10.1136/emmermed-2018-207570)] [Medline: [29921622](https://pubmed.ncbi.nlm.nih.gov/29921622/)]
19. Reed MJ, Grubb NR, Lang CC, O'Brien R, Simpson K, Padarenga M, et al. Multi-centre randomised controlled trial of a smartphone-based event recorder alongside standard care versus standard care for patients presenting to the emergency department with palpitations and pre-syncope: the IPED (Investigation of Palpitations in the ED) study. *EClinicalMedicine* 2019 Feb;8:37-46 [FREE Full text] [doi: [10.1016/j.eclinm.2019.02.005](https://doi.org/10.1016/j.eclinm.2019.02.005)] [Medline: [31193636](https://pubmed.ncbi.nlm.nih.gov/31193636/)]
20. Valenzuela C. [2 solutions for estimating odds ratios with zeros]. *Rev Med Chil* 1993 Dec;121(12):1441-1444. [Medline: [8085071](https://pubmed.ncbi.nlm.nih.gov/8085071/)]
21. Harrison S, Jones HE, Martin RM, Lewis SJ, Higgins JPT. The albatross plot: a novel graphical tool for presenting results of diversely reported studies in a systematic review. *Res Synth Methods* 2017 Sep;8(3):281-289 [FREE Full text] [doi: [10.1002/jrsm.1239](https://doi.org/10.1002/jrsm.1239)] [Medline: [28453179](https://pubmed.ncbi.nlm.nih.gov/28453179/)]
22. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
23. Liu J, Fang P, Hou Y, Li X, Liu Y, Wang Y, et al. The value of transtelephonic electrocardiogram monitoring system during the "Blanking Period" after ablation of atrial fibrillation. *J Electrocardiol* 2010;43(6):667-672. [doi: [10.1016/j.jelectrocard.2010.06.007](https://doi.org/10.1016/j.jelectrocard.2010.06.007)] [Medline: [20667550](https://pubmed.ncbi.nlm.nih.gov/20667550/)]
24. Rosenberg MA, Samuel M, Thosani A, Zimetbaum PJ. Use of a noninvasive continuous monitoring device in the management of atrial fibrillation: a pilot study. *Pacing Clin Electrophysiol* 2013 Mar;36(3):328-333 [FREE Full text] [doi: [10.1111/pace.12053](https://doi.org/10.1111/pace.12053)] [Medline: [23240827](https://pubmed.ncbi.nlm.nih.gov/23240827/)]
25. 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 Jan;127(1):e11-e17 [FREE Full text] [doi: [10.1016/j.amjmed.2013.10.003](https://doi.org/10.1016/j.amjmed.2013.10.003)] [Medline: [24384108](https://pubmed.ncbi.nlm.nih.gov/24384108/)]
26. Hendrikx T, Rosenqvist M, Wester P, Sandström H, Hörnsten R. Intermittent short ECG recording is more effective than 24-hour Holter ECG in detection of arrhythmias. *BMC Cardiovasc Disord* 2014 Apr 01;14:41 [FREE Full text] [doi: [10.1186/1471-2261-14-41](https://doi.org/10.1186/1471-2261-14-41)] [Medline: [24690488](https://pubmed.ncbi.nlm.nih.gov/24690488/)]
27. Kimura T, Aizawa Y, Kurata N, Nakajima K, Kashimura S, Kunitomi A, et al. Assessment of atrial fibrillation ablation outcomes with clinic ECG, monthly 24-h Holter ECG, and twice-daily telemonitoring ECG. *Heart Vessels* 2017 Mar;32(3):317-325. [doi: [10.1007/s00380-016-0866-2](https://doi.org/10.1007/s00380-016-0866-2)] [Medline: [27385021](https://pubmed.ncbi.nlm.nih.gov/27385021/)]
28. Busch MC, Gross S, Alte D, Kors JA, Völzke H, Ittermann T, et al. Impact of atrial fibrillation detected by extended monitoring: a population-based cohort study. *Ann Noninvasive Electrocardiol* 2017 Nov;22(6):1. [doi: [10.1111/anec.12453](https://doi.org/10.1111/anec.12453)] [Medline: [28440600](https://pubmed.ncbi.nlm.nih.gov/28440600/)]
29. Halcox JPJ, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, et al. Assessment of remote heart rhythm sampling using the alivecor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study. *Circulation* 2017 Nov 07;136(19):1784-1794. [doi: [10.1161/CIRCULATIONAHA.117.030583](https://doi.org/10.1161/CIRCULATIONAHA.117.030583)] [Medline: [28851729](https://pubmed.ncbi.nlm.nih.gov/28851729/)]
30. Hickey K, Biviano A, Garan H, Sciacca RR, Riga T, Warren K, et al. Evaluating the utility of mHealth ECG heart monitoring for the detection and management of atrial fibrillation in clinical practice. *J Atr Fibrillation* 2017;9(5):1546 [FREE Full text] [doi: [10.4022/jafib.1546](https://doi.org/10.4022/jafib.1546)] [Medline: [29250277](https://pubmed.ncbi.nlm.nih.gov/29250277/)]
31. Narasimha D, Hanna N, Beck H, Chaskes M, Glover R, Gatewood R, et al. Validation of a smartphone-based event recorder for arrhythmia detection. *Pacing Clin Electrophysiol* 2018 May;41(5):487-494. [doi: [10.1111/pace.13317](https://doi.org/10.1111/pace.13317)] [Medline: [29493801](https://pubmed.ncbi.nlm.nih.gov/29493801/)]
32. Goldenthal IL, Sciacca RR, Riga T, Bakken S, Baumeister M, Biviano AB, et al. Recurrent atrial fibrillation/flutter detection after ablation or cardioversion using the AliveCor KardiaMobile device: iHEART results. *J Cardiovasc Electrophysiol* 2019 Nov;30(11):2220-2228 [FREE Full text] [doi: [10.1111/jce.14160](https://doi.org/10.1111/jce.14160)] [Medline: [31507001](https://pubmed.ncbi.nlm.nih.gov/31507001/)]
33. Karunadas CP, Mathew C. Comparison of arrhythmia detection by conventional Holter and a novel ambulatory ECG system using patch and Android app over 24 hour period. *Indian Pacing Electrophysiol J* 2020;20(2):49-53 [FREE Full text] [doi: [10.1016/j.ipej.2019.12.013](https://doi.org/10.1016/j.ipej.2019.12.013)] [Medline: [31866554](https://pubmed.ncbi.nlm.nih.gov/31866554/)]
34. Kaura A, Szturha 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 Jul 26;24(1):25 [FREE Full text] [doi: [10.1186/s40001-019-0383-8](https://doi.org/10.1186/s40001-019-0383-8)] [Medline: [31349792](https://pubmed.ncbi.nlm.nih.gov/31349792/)]
35. AliveCor. URL: <https://www.alivecor.com/> [accessed 2020-11-17]
36. Zenicor. URL: <https://www.zenicor.com/> [accessed 2020-11-17]
37. Jaakkola J, Mustonen P, Kiviniemi T, Hartikainen JEK, Palomäki A, Hartikainen P, et al. Stroke as the first manifestation of atrial fibrillation. *PLoS One* 2016;11(12):e0168010 [FREE Full text] [doi: [10.1371/journal.pone.0168010](https://doi.org/10.1371/journal.pone.0168010)] [Medline: [27936187](https://pubmed.ncbi.nlm.nih.gov/27936187/)]

38. Welton NJ, McAleenan A, Thom HH, Davies P, Hollingworth W, Higgins JP, et al. Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2017 May;21(29):1-236 [FREE Full text] [doi: [10.3310/hta21290](https://doi.org/10.3310/hta21290)] [Medline: [28629510](https://pubmed.ncbi.nlm.nih.gov/28629510/)]
39. Lowres N, Olivier J, Chao T, Chen S, Chen Y, Diederichsen A, et al. Estimated stroke risk, yield, and number needed to screen for atrial fibrillation detected through single time screening: a multicountry patient-level meta-analysis of 141,220 screened individuals. *PLoS Med* 2019 Sep;16(9):e1002903 [FREE Full text] [doi: [10.1371/journal.pmed.1002903](https://doi.org/10.1371/journal.pmed.1002903)] [Medline: [31553733](https://pubmed.ncbi.nlm.nih.gov/31553733/)]
40. Giebel GD. Use of mHealth devices to screen for atrial fibrillation: cost-effectiveness analysis. *JMIR Mhealth Uhealth* 2020 Oct 06;8(10):e20496 [FREE Full text] [doi: [10.2196/20496](https://doi.org/10.2196/20496)] [Medline: [33021489](https://pubmed.ncbi.nlm.nih.gov/33021489/)]
41. Bacchini M, Bonometti S, Del Zotti F, Lechi A, Realdon F, Fava C, et al. Opportunistic screening for atrial fibrillation in the pharmacies: a population-based cross-sectional study. *High Blood Press Cardiovasc Prev* 2019 Aug;26(4):339-344. [doi: [10.1007/s40292-019-00334-4](https://doi.org/10.1007/s40292-019-00334-4)] [Medline: [31385256](https://pubmed.ncbi.nlm.nih.gov/31385256/)]
42. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies: the SEARCH-AF study. *Thromb Haemost* 2014 Jun;111(6):1167-1176. [doi: [10.1160/TH14-03-0231](https://doi.org/10.1160/TH14-03-0231)] [Medline: [24687081](https://pubmed.ncbi.nlm.nih.gov/24687081/)]
43. Lowres N, Krass I, Neubeck L, Redfern J, McLachlan AJ, Bennett AA, et al. Atrial fibrillation screening in pharmacies using an iPhone ECG: a qualitative review of implementation. *Int J Clin Pharm* 2015 Dec;37(6):1111-1120. [doi: [10.1007/s11096-015-0169-1](https://doi.org/10.1007/s11096-015-0169-1)] [Medline: [26202627](https://pubmed.ncbi.nlm.nih.gov/26202627/)]
44. Kaasenbrood F, Hollander M, Rutten FH, Gerhards LJ, Hoes AW, Tieleman RG. Yield of screening for atrial fibrillation in primary care with a hand-held, single-lead electrocardiogram device during influenza vaccination. *Europace* 2016 Oct;18(10):1514-1520 [FREE Full text] [doi: [10.1093/europace/euv426](https://doi.org/10.1093/europace/euv426)] [Medline: [26851813](https://pubmed.ncbi.nlm.nih.gov/26851813/)]
45. Orchard J, Lowres N, Freedman SB, Ladak L, Lee W, Zwar N, et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): a feasibility study. *Eur J Prev Cardiol* 2016 Dec;23(2 suppl):13-20. [doi: [10.1177/2047487316670255](https://doi.org/10.1177/2047487316670255)] [Medline: [27892421](https://pubmed.ncbi.nlm.nih.gov/27892421/)]
46. Calkins H, Brugada J, Packer DL, Cappato R, Chen S, Crijns HJG, Heart Rhythm Society, European Heart Rhythm Association, European Cardiac Arrhythmia Society, American College of Cardiology, American Heart Association, Society of Thoracic Surgeons. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for personnel, policy, procedures and follow-up. A report of the Heart Rhythm Society (HRS) Task Force on Catheter and Surgical Ablation of Atrial Fibrillation developed in partnership with the European Heart Rhythm Association (EHRA) and the European Cardiac Arrhythmia Society (ECAS); in collaboration with the American College of Cardiology (ACC), American Heart Association (AHA), and the Society of Thoracic Surgeons (STS). Endorsed and approved by the governing bodies of the American College of Cardiology, the American Heart Association, the European Cardiac Arrhythmia Society, the European Heart Rhythm Association, the Society of Thoracic Surgeons, and the Heart Rhythm Society. *Europace* 2007 Jun;9(6):335-379. [doi: [10.1093/europace/eum120](https://doi.org/10.1093/europace/eum120)] [Medline: [17599941](https://pubmed.ncbi.nlm.nih.gov/17599941/)]
47. Mantovan R, Macle L, De Martino G, Chen J, Morillo CA, Novak P, et al. Relationship of quality of life with procedural success of atrial fibrillation (AF) ablation and postablation AF burden: substudy of the STAR AF randomized trial. *Can J Cardiol* 2013 Oct;29(10):1211-1217. [doi: [10.1016/j.cjca.2013.06.006](https://doi.org/10.1016/j.cjca.2013.06.006)] [Medline: [23988341](https://pubmed.ncbi.nlm.nih.gov/23988341/)]
48. Boriani G, Glotzer TV, Santini M, West TM, De Melis M, Sepsi M, et al. Device-detected atrial fibrillation and risk for stroke: an analysis of >10,000 patients from the SOS AF project (Stroke preventiOn Strategies based on Atrial Fibrillation information from implanted devices). *Eur Heart J* 2014 Feb;35(8):508-516 [FREE Full text] [doi: [10.1093/eurheartj/eh491](https://doi.org/10.1093/eurheartj/eh491)] [Medline: [24334432](https://pubmed.ncbi.nlm.nih.gov/24334432/)]
49. Steinberg JS, O'Connell H, Li S, Ziegler PD. Thirty-second gold standard definition of atrial fibrillation and its relationship with subsequent arrhythmia patterns. *Circ Arrhythm Electrophysiol* 2018 Jul;11(7):e006274. [doi: [10.1161/circep.118.006274](https://doi.org/10.1161/circep.118.006274)]
50. Apple Inc. URL: <https://support.apple.com/en-us/HT208955> [accessed 2020-11-17]
51. Withings. URL: <https://www.withings.com/> [accessed 2020-11-17]
52. Risling T, Martinez J, Young J, Thorp-Froslic N. Evaluating patient empowerment in association with eHealth technology: scoping review. *J Med Internet Res* 2017 Sep 29;19(9):e329 [FREE Full text] [doi: [10.2196/jmir.7809](https://doi.org/10.2196/jmir.7809)] [Medline: [28963090](https://pubmed.ncbi.nlm.nih.gov/28963090/)]

## Abbreviations

**AF:** atrial fibrillation

**CHA<sub>2</sub>DS<sub>2</sub>-VASc:** congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke or transient ischemic attack (TIA), vascular disease, age 65 to 74 years, sex category

**ECG:** electrocardiogram

**EPACS:** Early Prolonged Ambulatory Cardiac Monitoring in Stroke

**iHeart:** An Information Technology Approach to Implementing Depression Treatment in Cardiac Patients

**IPED:** Investigation of Palpitations in the Emergency Department

**mHealth:** mobile health

**OR:** odds ratio

**PAC:** premature atrial contraction

**PATCH-ED:** Patch Monitor in Patients With Unexplained Syncope After Initial Evaluation in the Emergency Department

**RCT:** randomized controlled trial

**REHEARSE-AF:** Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation

**RoB 2:** Risk of Bias 2

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

*Edited by L Buis; submitted 30.11.20; peer-reviewed by KL Mauco, G Giebel; comments to author 01.02.21; revised version received 25.02.21; accepted 22.03.21; published 28.04.21.*

*Please cite as:*

*Biersteker TE, SchaliJ MJ, Treskes RW*

*Impact of Mobile Health Devices for the Detection of Atrial Fibrillation: Systematic Review*

*JMIR Mhealth Uhealth 2021;9(4):e26161*

URL: <https://mhealth.jmir.org/2021/4/e26161>

doi: [10.2196/26161](https://doi.org/10.2196/26161)

PMID: [33908885](https://pubmed.ncbi.nlm.nih.gov/33908885/)

©Tom E Biersteker, Martin J SchaliJ, Roderick W Treskes. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 28.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Accuracy of Smart Scales on Weight and Body Composition: Observational Study

Justine Frija-Masson<sup>1,2,3</sup>, MD, DPhil; Jimmy Mullaert<sup>4,5</sup>, MD, DPhil; Emmanuelle Vidal-Petiot<sup>1,6</sup>, MD, DPhil; Nathalie Pons-Kerjean<sup>3,7</sup>, MPH, DPhil; Martin Flamant<sup>1,6</sup>, MD, DPhil; Marie-Pia d'Ortho<sup>1,2,3</sup>, MD, DPhil

<sup>1</sup>Physiologie-Explorations Fonctionnelles, Fédération Hospitalo-Universitaire APOLLO (Personalised medicine in chronic cardiovascular, respiratory, renal diseases and organ transplantation), Hôpital Bichat Claude Bernard, Assistance Publique Hôpitaux de Paris, Paris, France

<sup>2</sup>Université de Paris, Neurodiderot, Institut national de la santé et de la recherche médicale U 1141, Paris, France

<sup>3</sup>Digital Medical Hub, Hôpital Bichat Claude Bernard, Assistance Publique Hôpitaux de Paris, Paris, France

<sup>4</sup>Département d'Epidémiologie, Biostatistiques et Recherche Clinique, Hôpital Bichat Claude Bernard, Assistance Publique Hôpitaux de Paris, Paris, France

<sup>5</sup>Université de Paris, Infection, antimicrobiens, modélisation, évolution, institut national de la santé et de la recherche médicale, Paris, France

<sup>6</sup>Université de Paris, Institut national de la santé et de la recherche médicale U 1149, Paris, France

<sup>7</sup>Pharmacie, Hôpital Bichat Claude Bernard, Assistance Publique Hôpitaux de Paris, Paris, France

**Corresponding Author:**

Justine Frija-Masson, MD, DPhil

Physiologie-Explorations Fonctionnelles, Fédération Hospitalo-Universitaire APOLLO (Personalised medicine in chronic cardiovascular, respiratory, renal diseases and organ transplantation)

Hôpital Bichat Claude Bernard

Assistance Publique Hôpitaux de Paris

46 rue Henri Huchard

Paris

France

Phone: 33 01 40 25 85 18

Email: [justine.frija@aphp.fr](mailto:justine.frija@aphp.fr)

## Abstract

**Background:** Smart scales are increasingly used at home by patients to monitor their body weight and body composition, but scale accuracy has not often been documented.

**Objective:** The goal of the research was to determine the accuracy of 3 commercially available smart scales for weight and body composition compared with dual x-ray absorptiometry (DEXA) as the gold standard.

**Methods:** We designed a cross-sectional study in consecutive patients evaluated for DEXA in a physiology unit in a tertiary hospital in France. There were no exclusion criteria except patient declining to participate. Patients were weighed with one smart scale immediately after DEXA. Three scales were compared (scale 1: Body Partner [Téfal], scale 2: DietPack [Terraillon], and scale 3: Body Cardio [Nokia Withings]). We determined absolute error between the gold standard values obtained from DEXA and the smart scales for body mass, fat mass, and lean mass.

**Results:** The sample for analysis included 53, 52, and 48 patients for each of the 3 tested smart scales, respectively. The median absolute error for body weight was 0.3 kg (interquartile range [IQR] -0.1, 0.7), 0 kg (IQR -0.4, 0.3), and 0.25 kg (IQR -0.10, 0.52), respectively. For fat mass, absolute errors were -2.2 kg (IQR -5.8, 1.3), -4.4 kg (IQR -6.6, 0), and -3.7 kg (IQR -8.0, 0.28), respectively. For muscular mass, absolute errors were -2.2 kg (IQR -5.8, 1.3), -4.4 kg (IQR -6.6, 0), and -3.65 kg (IQR -8.03, 0.28), respectively. Factors associated with fat mass measurement error were weight for scales 1 and 2 ( $P=.03$  and  $P<.001$ , respectively), BMI for scales 1 and 2 ( $P=.034$  and  $P<.001$ , respectively), body fat for scale 1 ( $P<.001$ ), and muscular and bone mass for scale 2 ( $P<.001$  for both). Factors associated with muscular mass error were weight and BMI for scale 1 ( $P<.001$  and  $P=.004$ , respectively), body fat for scales 1 and 2 ( $P<.001$  for both), and muscular and bone mass for scale 2 ( $P<.001$  and  $P=.002$ , respectively).

**Conclusions:** Smart scales are not accurate for body composition and should not replace DEXA in patient care.

**Trial Registration:** ClinicalTrials.gov NCT03803098; <https://clinicaltrials.gov/ct2/show/NCT03803098>

**KEYWORDS**

smart scales; DEXA; obesity

## Introduction

Cellular-connected scales, familiarly known as smart scales, are increasingly used at home for weight follow-up. They have been shown to increase the frequency of self-weighing and weight loss [1,2]. They can be connected to other smart objects, such as motion sensors, and thus may help subjects engage in greater physical activity and better nutritional habits. Most available smart scales combine a classic weight scale with a foot-to-foot impedance meter (FFI) that can estimate body composition (ie, fat mass [FM] and fat-free mass [FFM]) by measuring foot-to-foot impedance at different frequency. Whole body FFM is calculated from a model comprising body impedance, height, weight, and age trained with dual x-ray absorptiometry (DEXA) data [3]. Smart scales are easier to use than medical impedance meters since they do not require a supine position and their electrodes are indefinitely reusable. But the accuracy of smart scales depends on the representativeness of the patient population used to train the model and the model itself. Although some FFIs have been compared with DEXA and to medical impedance meters [3,4], no data are available for smart scale FFIs, and the regression equations used are unknown. The purpose of this cross-sectional study was to assess metrologic accuracy of 3 commercially available smart scales compared with DEXA as the gold standard in a population of adult patients from a tertiary hospital physiology unit.

## Methods

### Patients

Consecutive adult patients evaluated for body composition by DEXA at the Bichat Hospital during the study were eligible. Patients were referred for DEXA because of obesity or a chronic condition that can affect body composition (eg, chronic kidney disease, long-term steroids). All patients were included in the analysis except those who refused to participate or whose weight measured by Lunar iDXA (GE Healthcare) exceeded the maximum weight tolerated by the smart scale (eg, 160 kg for DietPack and 180 kg for Body Partner and Body Cardio). Weight and body composition by DEXA was performed by a trained technician according to current practice on a Lunar

iDXA. Three smart scales were assessed (Body Partner [Téfal], DietPack [Terrailon], and Body Cardio [Nokia Withings]) during 3 consecutive periods (ie, only 1 scale was tested per patient for feasibility reasons). Patients were not required to be fasting. DEXA and assessment by the smart scale were performed with the patient wearing a hospital disposable gown and their undergarment. A total of 55 patients were planned for each scale.

### Statistics

Patient characteristics were described as median and interquartile range for quantitative variables and percentages for categorical variables. Absolute and relative errors for each scale with respect to DEXA were reported with median and interquartile range. Bland-Altman representations were used to show systematic bias or trend in measurement error. Univariate linear models were used to estimate and test the association between measurement error and possible associated variables. For this analysis, we reported the slope estimation with its 95% confidence interval and the *P* value corresponding to the Wald statistic. We performed no imputation for missing data. The significance threshold was .05. All analyses were performed using R software version 4.0.2 (R Foundation for Statistical Computing).

### Ethical Aspects

This study is part of the Evaluation of the Metrological Reliability of Connected Objects in the Measurement of Medical Physiological Parameters (EvalExplo) study [NCT03803098]. Ethics approval was obtained from Comité de Protection des Personnes Sud Est VI (approval number AU 1443), and written nonopposition was obtained.

## Results

### Patient Characteristics

The final sample for analysis included 53, 52, and 48 patients for each scale, respectively, after taking into account missing data (eg, smart scale not retrieving data despite several attempts for all but one who was excluded because of excessive weight for the scales). Patient characteristics are presented in [Table 1](#). There were no significant differences between the 3 groups.

**Table 1.** Description of included patients.

Characteristic	Scale 1: Body Partner (n=53)	Scale 2: DietPack (n=52)	Scale 3: Body Cardio (n=48)
Sex, male, n (%)	17 (32)	27 (52)	32 (67)
Age in years, median (IQR <sup>a</sup> )	46 (37-58)	48 (40-56)	51 (43-62)
Weight (kg), median (IQR)	83.4 (71.4-102.1)	80.5 (69.2-99.3)	82.2 (75-94)
Height (cm), median (IQR)	166 (162-173)	169.5 (164-175)	166 (159-173)
<b>BMI (kg/m<sup>2</sup>), median (IQR)</b>	30.3 (25.0-37.6)	28.4 (24.8-34.1)	31.2 (26.1-35.5)
BMI < 25, n (%)	13 (25)	14 (27)	9 (19)
25 ≤ BMI < 30, n (%)	11 (21)	16 (31)	13 (27)
30 ≤ BMI < 35, n (%)	14 (26)	11 (21)	12 (25)
BMI ≥ 35, n (%)	15 (28)	11 (21)	14 (29)
Body fat (kg), median (IQR)	33.5 (21.5-49.8)	31.2 (23.5-43.4)	28.5 (16.0-41.4)
Muscular mass (kg), median (IQR)	49.1 (42.0-56.1)	52.0 (43.8-57.1)	47.6 (42.8-53.8)
Bone mass (kg), median (IQR)	2.77 (2.23-3.02)	2.74 (2.33-3.21)	2.56 (2.17-2.93)
Excess mass (kg), median (IQR)	1.04 (0.84-1.29)	0.98 (0.81-1.27)	4.8 (2.2-6.2)

<sup>a</sup>IQR: interquartile range.

### Accuracy of the Scales

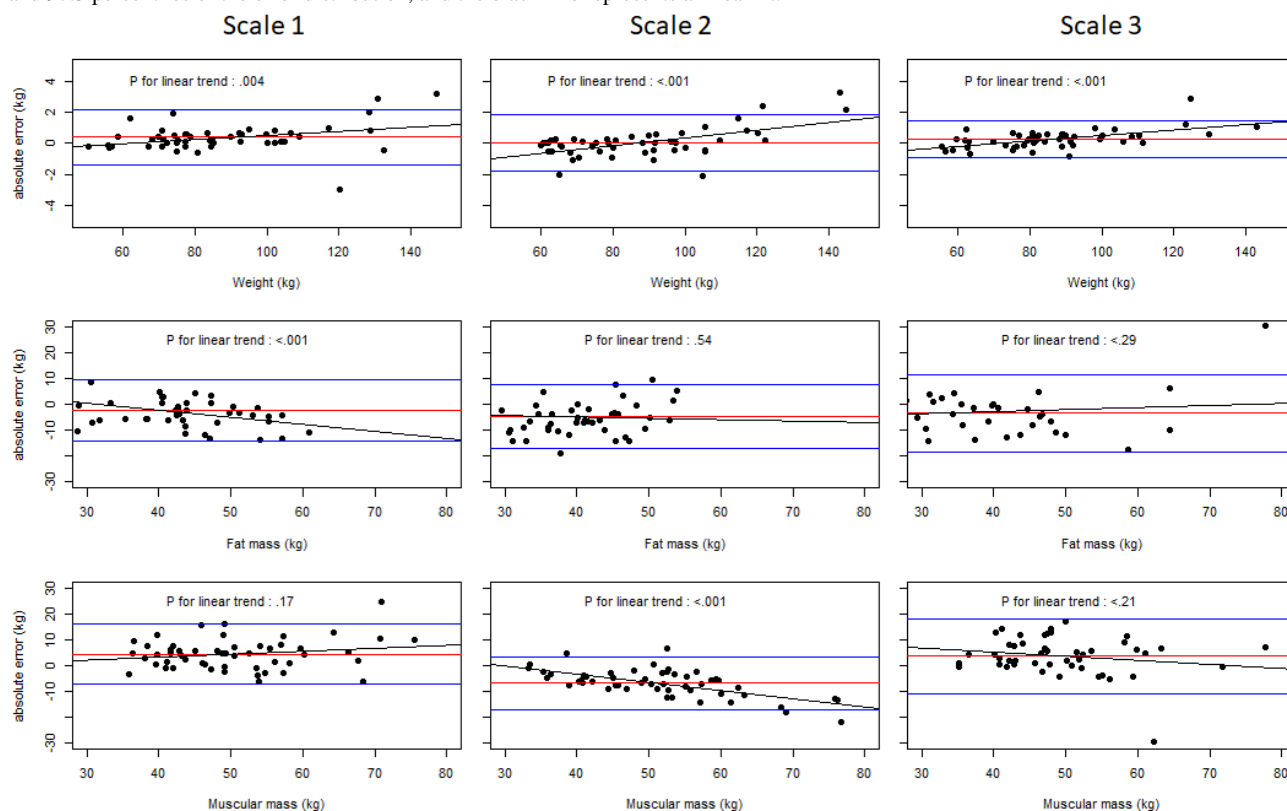
All 3 scales gave rather accurate weights, with a median absolute difference of less than a kilogram compared with DEXA. The median absolute error for body weight was 0.3 kg (interquartile range [IQR] -0.1, 0.7), 0 kg (IQR -0.4, 0.3), and 0.25 kg (IQR -0.1, 0.5) for Body Partner (scale 1), DietPack (scale 2), and Body Cardio (scale 3), respectively.

For fat mass, absolute errors were -2.2 kg (IQR -5.8, 1.3), -4.4 kg (IQR -6.6, 0), and -3.7 kg (IQR -8.0, 0.3), respectively. Body fat was globally underestimated in all 3 scales. For

muscular mass, absolute errors were 4.50 kg (IQR 0.4, 7.3), -6.6 kg (IQR -9.4, -3.6), and 4.0 kg (IQR 0.1, 7.6), respectively. Muscular mass was thus underestimated by scale 2 and overestimated by scales 1 and 3.

Bland-Altman graphs are presented in [Figure 1](#) for weight, body fat, and muscular mass for the 3 scales. They show a significant linear trend of body weight on measured weight, but absolute errors remain reasonable and compatible with clinical practice. However, significant linear trends for fat and muscular mass (scales 1 and 2, respectively) lead to very high errors.

**Figure 1.** Bland-Altman plots of the three scales for body weight, fat mass, and muscular mass. The red line indicates mean error, blue lines indicate 2.5 and 97.5 percentiles of the error distribution, and the black line represents a linear fit.



### Factors Associated With Measurement Error

Tables 2 and 3 show the factors associated with fat and muscular mass measurement error. Factors associated with fat mass measurement error were weight for scales 1 and 2 ( $P=.03$  and  $P<.001$ , respectively), BMI for scales 1 and 2 ( $P=.034$  and  $P<.001$ , respectively), body fat for scale 1 ( $P<.001$ ), and muscular and bone mass for scale 2 ( $P<.001$  and  $P<.001$ ,

respectively). Factors associated with muscular mass error were weight and BMI for scale 1 ( $P<.001$  and  $P=.004$ , respectively), body fat for scales 1 and 2 ( $P<.001$  and  $P<.001$ , respectively), and muscular and bone mass for scale 2 ( $P<.001$  and  $P=.002$ , respectively).

We found no factor associated with measurement error of fat or muscular mass for scale 3. Sex did not show a significant influence on measurement error.



**Table 2.** Factors associated with fat mass measurement error for the 3 scales (univariate linear regressions).

Characteristic	Scale 1: Body Partner (n=53)			Scale 2: DietPack (n=52)			Scale 3: Body Cardio (n=48)		
	Estimate	95% CI	P value	Estimate	95% CI	P value	Estimate	95% CI	P value
<b>Sex</b>	— <sup>a</sup>	—	.96	—	—	.24	—	—	.57
Female	1 (ref)	—	—	1 (ref)	—	—	1 (ref)	—	—
Male	0.095	-3.5 to 3.7	—	2.1	-1.4 to 5.7	—	1.3	-3.4 to 6.0	—
Age in years	0.023	-0.091 to 0.14	.69	0.042	-0.088 to 0.17	.52	-0.064	-0.25 to 0.12	.49
Weight (kg)	-0.065	-0.12 to -0.0	.03	0.18	0.11 to 0.24	<.001	0.079	-0.032 to 0.190	.16
Height (cm)	0.024	-0.15 to 0.20	.79	0.054	-0.17 to 0.28	.64	0.2	-0.035 to 0.43	.09
<b>BMI (kg/m<sup>2</sup>)</b>	—	—	.03	—	—	<.001	—	—	.83
BMI<25	1 (ref)	—	—	1 (ref)	—	—	1 (ref)	—	—
25<BMI<30	2	-2.8 to 6.7	—	1.2	-2.8 to 5.2	—	0.97	-5.8 to -8.4	—
30<BMI<35	-2.8	-7.2 to 1.6	—	3.8	-0.55 to 8.2	—	-1.5	-8.4 to 5.4	—
BMI>35	-4.4	-8.7 to -0.013	—	9.8	5.42 to 14.2	—	-1.4	-8.0 to 5.3	—
Body fat (kg)	-0.27	-0.40 to -0.14	<.001	-0.053	-0.23 to 0.12	.54	0.08	-0.07 to 0.23	.29
Muscular mass (kg)	0.0081	-0.16 to 0.17	.92	0.42	0.29 to 0.54	<.001	0.16	-0.084 to 0.41	.19
Bone mass (kg)	-0.45	-3.4 to 2.5	.76	6.2	3.2 to 9.3	<.001	2.4	-1.6 to 6.3	.24

<sup>a</sup>Not applicable.

**Table 3.** Factors associated with muscular mass measurement error for the 3 scales (univariate linear regressions).

Characteristic	Scale 1: Body Partner (n=53)			Scale 2: DietPack (n=52)			Scale 3: Body Cardio (n=48)		
	Estimate	95% CI	P value	Estimate	95% CI	P value	Estimate	95% CI	P value
<b>Sex</b>	— <sup>a</sup>	—	.52	—	—	.07	—	—	.47
Male	1 (ref)	—	—	1 (ref)	—	—	1 (ref)	—	—
Female	1.1	-2.4 to 4.7	—	-2.6	-5.47 to 0.26	—	-1.7	-6.3 to 2.9	—
Age in years	-0.039	-0.150 to 0.072	.49	0.098	-0.0054 to 0.20	.06	0.073	-0.11 to 0.25	.42
Weight (kg)	0.11	0.054 to 0.16	<.001	-0.048	-0.12 to 0.02	.16	-0.068	-0.18 to 0.042	.22
Height (cm)	0.085	-0.087 to 0.26	.33	-0.15	-0.331 to 0.032	.10	-0.2	-0.43 to 0.025	.08
<b>BMI (kg/m<sup>2</sup>)</b>	—	—	.003	—	—	.87	—	—	.77
BMI<25	1 (ref)	—	—	1 (ref)	—	—	1 (ref)	—	—
25<BMI<30	-1.9	-6.3 to 2.5	—	-0.43	-4.4 to 3.5	—	-0.41	-7.0 to 6.2	—
30<BMI<35	1.9	-2.2 to 6.0	—	-1.67	-6.0 to 2.7	—	2	-4.7 to 8.7	—
BMI>35	5.8	1.8 to 9.9	—	-1.11	-5.5 to 3.2	—	2.14	-4.4 to 8.7	—
Body fat (kg)	0.28	0.15 to 0.40	<.001	0.23	0.10 to 0.36	<.001	-0.063	-0.21 to 0.085	.40
Muscular mass (kg)	0.11	-0.047 to 0.27	.17	-0.32	-0.42 to -0.21	<.001	-0.15	-0.40 to 0.089	.21
Bone mass (kg)	2.2	-0.64 to 4.94	.13	-4.2	-6.9 to -1.6	.002	-2.1	-6.0 to 1.9	.30

<sup>a</sup>Not applicable.

## Discussion

### Principal Findings

To our knowledge, this is the first study to assess metrologic accuracy of commercially available smart scales (scale 1: Body Partner, scale 2: DietPack, scale 3: Body Cardio). We show that all scales were reasonably accurate for body weight but not body composition. Total body weight was associated with fat

mass and muscular measurement error for scales 1 and 2, but we were not able to find factors associated with error measurement for scale 3.

### Possible Explanations for the Lack of Accuracy

Smart scales combine a classic weight scale and an FFI, which has been widely available for almost three decades and has been compared with medical impedance meters and DEXA in several publications [5-8]. They have been shown to be more sensitive

to differences in morphology than whole body impedance measurements, since their data depend upon an extrapolation of measurements made on the lower part of the body to the entire body. Indeed, Bousbiat et al [3] conducted an extensive technical study on FFIs. They found that measurement can be affected by surface contact (ie, foot size and width) and sweat but also by foot position on the scale and flexion of the legs. Surface contact with the electrodes will vary depending on the subject's foot length and width, and thus can be affected by total body weight and total height. Since there is no precise guidance on the scales for subjects on where to put their feet during body composition estimation, this may partly explain the differences between DEXA and scales. In clinical settings, clinicians or technicians should pay attention to the subject's position during measurement. In the same way, subjects should be advised not to bend their legs. At home, subjects should try to follow directions given by the scale as closely as possible and keep the same position on the scale for follow-up.

In FFIs, whole body FFM is calculated from a regression equation expressing resistance generally as a function of height, weight, and age determined by comparison with DEXA data, while FM is calculated as the weight of FFM. Each smart scale has its own regression equation. It is thus plausible that BMI can affect error measurement. We did not find any factor associated with error measurement for scale 3 despite a very high dispersion, which can have different explanations: our study may be underpowered to detect such association or unobserved variables not part of the secret regression implemented in the smart scale may explain the residual error.

### Clinical Relevance of These Data

Weight was accurately measured by all 3 scales, but body fat was underestimated. For scales 1 and 2, we found a significant effect of higher body weight on fat mass error; this error remains small compared with total body weight in patients with obesity but can be of importance in patients with normal or underweight. Ross et al [9] compared the weight from in-person visits and BodyTrace brand smart scales in 58 patients and found a mean bias of 1.1 (SD 0.8) kg, 95% CI 0.5 to 2.6, but agreement seemed lower for patients weighing above 110 kg.

It is unlikely that body composition will be followed by DEXA in the clinical setting, since DEXA uses x-rays. Since smart scales are widely available, it is possible that patients will follow their body composition at home. Thus, if the first body composition is assessed by DEXA, clinicians and patients should be aware that there might be a difference in body composition, which can reach 1 kg. Also, follow-up in the clinical setting should use the same connected device to ensure repeatability of measurements.

### Interest for Follow-Up

Despite very poor accuracy for estimating body composition, some authors reported potential interests of in-home use of smart scales. Indeed, in several randomized studies, when

compared with commercial weight management programs or standard weight loss counseling, smart scale use was found to allow a greater proportion of participants to achieve significant weight loss after several months (3 to 12 months) [1,2,10]. In these studies, no data were available on body composition and its evolution. Also, several studies report a greater weight loss in patients using smart scales than in patients self-reporting weight loss or being weighted only during visits [11,12]. Last, data exist on the importance of weight variability in final weight loss and weight maintenance [13-15]; using smart scales at home with automated data treatment could help health care professionals and patients in achieving and maintaining greater weight loss [16].

### Strengths and Limitations

To our knowledge, this is the first cross-sectional study assessing metrologic accuracy of commercially available smart scales. We compared these scales to DEXA, which is considered the gold standard for body composition assessment.

Included patients were evaluated for body composition either because of obesity or because of a chronic condition (steroids, chronic renal failure, etc) without obesity. Thus, our sample covers a wide range of body weight: underweight, lean, and obese patients. The setting of the study in a tertiary hospital explains the relatively small proportion of patients with obesity since the main reason for prescribing DEXA was long-term treatment with steroids. A specific study focusing on patients with obesity and extreme obesity would be required since they are likely to benefit the most from body composition evaluation and follow-up and since the error on fat mass is affected by total weight. This should be done keeping in mind the maximum weight supported by the scales (180 kg for Body Cardio and Body Partner and 160 kg for DietPack), while DEXA supports higher weight (230 kg on our machine).

Finally, although body composition estimation is quick with the smart scales, total experimental time was 15 to 30 minutes for each scale due to scale and app setup and patient information. Thus, it was not possible to evaluate the 3 scales on the same patient or replicate measures. However, due to the very poor accuracy of these scales shown in this study, comparisons between scales and estimation of intraindividual variability would have limited relevance. This study included only a limited number of patients. However, while it is always possible that the studied scales would perform better on an independent sample, the study results are so clear that it is also unlikely that a higher sample size could change our conclusion relative to the accuracy of body composition estimation.

### Conclusion

Our study shows that although smart scales are accurate for total body weight, they should not be used routinely to assess body composition, especially in patients with severe obesity. Further studies are needed to clarify their utility in patient follow-up.

## Acknowledgments

The authors are thankful to Najat Benlahmar and Fedja Kerzabi for data acquisition and Chanez Chemani, Caroline Quintin, and Camille Couffignal for helpful methodologic discussion. The EvalExplo study was funded through a donation from MSD France to the Assistance Publique Hôpitaux de Paris Foundation. The authors are thankful to François Crémieux for his support in creating the Digital Medical Hub.

## Conflicts of Interest

JFM has received travel, accommodation, and meeting expenses unrelated to this work from Vitalaire, LVL Medical, Oxyvie, Ixxair Medical, and Boehringer-Ingelheim. No conflicts were declared for other authors.

## References

1. Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. The efficacy of a daily self-weighing weight loss intervention using smart scales and e-mail. *Obesity (Silver Spring)* 2013 Sep;21(9):1789-1797 [FREE Full text] [doi: [10.1002/oby.20396](https://doi.org/10.1002/oby.20396)] [Medline: [23512320](https://pubmed.ncbi.nlm.nih.gov/23512320/)]
2. Thomas JG, Raynor HA, Bond DS, Luke AK, Cardoso CC, Wojtanowski AC, et al. Weight loss and frequency of body-weight self-monitoring in an online commercial weight management program with and without a cellular-connected 'smart' scale: a randomized pilot study. *Obes Sci Pract* 2017 Dec;3(4):365-372 [FREE Full text] [doi: [10.1002/osp4.132](https://doi.org/10.1002/osp4.132)] [Medline: [29259794](https://pubmed.ncbi.nlm.nih.gov/29259794/)]
3. Bousbiat S, Jaffrin M, Assadi I. Factors influencing accuracy and reproducibility of body resistance measurements by foot-to-foot impedancemeters. *J Med Eng Technol* 2015 Jan;39(1):35-43. [doi: [10.3109/03091902.2014.973617](https://doi.org/10.3109/03091902.2014.973617)] [Medline: [25365933](https://pubmed.ncbi.nlm.nih.gov/25365933/)]
4. Jaffrin MY, Moreno MV. Measurements of total body water with a foot-to-foot impedancemeter. *Med Eng Phys* 2008 May;30(4):483-489. [doi: [10.1016/j.medengphy.2007.05.009](https://doi.org/10.1016/j.medengphy.2007.05.009)] [Medline: [17629537](https://pubmed.ncbi.nlm.nih.gov/17629537/)]
5. Bousbiat S, Jaffrin MY, Dongmo E. Comparison of body fat-free masses calculated from hand-to-foot and foot-to-foot resistances with DXA measurements. *Med Biol Eng Comput* 2011 Nov;49(11):1329-1336. [doi: [10.1007/s11517-011-0827-y](https://doi.org/10.1007/s11517-011-0827-y)] [Medline: [21915720](https://pubmed.ncbi.nlm.nih.gov/21915720/)]
6. Jaffrin MY, Bousbiat S. Accuracy of plantar electrodes compared with hand and foot electrodes in fat-free-mass measurement. *J Healthc Eng* 2014;5(2):123-144 [FREE Full text] [doi: [10.1260/2040-2295.5.2.123](https://doi.org/10.1260/2040-2295.5.2.123)] [Medline: [24918180](https://pubmed.ncbi.nlm.nih.gov/24918180/)]
7. Jaffrin MY, Bousbiat S, Dongmo E. A comparison between two methods for measuring limb resistances with wrist and ankle electrodes. *Med Eng Phys* 2011 Oct;33(8):943-949. [doi: [10.1016/j.medengphy.2011.03.003](https://doi.org/10.1016/j.medengphy.2011.03.003)] [Medline: [21481627](https://pubmed.ncbi.nlm.nih.gov/21481627/)]
8. Jaffrin MY, Morel H. Measurements of body composition in limbs and trunk using a eight contact electrodes impedancemeter. *Med Eng Phys* 2009 Nov;31(9):1079-1086. [doi: [10.1016/j.medengphy.2009.07.005](https://doi.org/10.1016/j.medengphy.2009.07.005)] [Medline: [19656715](https://pubmed.ncbi.nlm.nih.gov/19656715/)]
9. Ross KM, Wing RR. Concordance of in-home "smart" scale measurement with body weight measured in-person. *Obes Sci Pract* 2016 Jun;2(2):224-248 [FREE Full text] [doi: [10.1002/osp4.41](https://doi.org/10.1002/osp4.41)] [Medline: [27499884](https://pubmed.ncbi.nlm.nih.gov/27499884/)]
10. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018 Dec;55(6):777-786 [FREE Full text] [doi: [10.1016/j.amepre.2018.07.005](https://doi.org/10.1016/j.amepre.2018.07.005)] [Medline: [30361140](https://pubmed.ncbi.nlm.nih.gov/30361140/)]
11. Sperrin M, Rushton H, Dixon WG, Normand A, Villard J, Chieh A, et al. Who self-weighs and what do they gain from it? A retrospective comparison between smart scale users and the general population in England. *J Med Internet Res* 2016 Jan 21;18(1):e17 [FREE Full text] [doi: [10.2196/jmir.4767](https://doi.org/10.2196/jmir.4767)] [Medline: [26794900](https://pubmed.ncbi.nlm.nih.gov/26794900/)]
12. Ross KM, Eastman A, Wing RR. Accuracy of self-report versus objective smart-scale weights during a 12-week weight management intervention. *Obesity (Silver Spring)* 2019 Mar;27(3):385-390. [doi: [10.1002/oby.22400](https://doi.org/10.1002/oby.22400)] [Medline: [30703282](https://pubmed.ncbi.nlm.nih.gov/30703282/)]
13. Benson L, Zhang F, Espel-Huynh H, Wilkinson L, Lowe MR. Weight variability during self-monitored weight loss predicts future weight loss outcome. *Int J Obes (Lond)* 2020 Jun;44(6):1360-1367. [doi: [10.1038/s41366-020-0534-6](https://doi.org/10.1038/s41366-020-0534-6)] [Medline: [31949298](https://pubmed.ncbi.nlm.nih.gov/31949298/)]
14. Turicchi J, O'Driscoll R, Lowe M, Finlayson G, Palmeira AL, Larsen SC, et al. The impact of early body-weight variability on long-term weight maintenance: exploratory results from the NoHoW weight-loss maintenance intervention. *Int J Obes (Lond)* 2021 Mar;45(3):525-534. [doi: [10.1038/s41366-020-00706-0](https://doi.org/10.1038/s41366-020-00706-0)] [Medline: [33144700](https://pubmed.ncbi.nlm.nih.gov/33144700/)]
15. Brockmann AN, Eastman A, Ross KM. Frequency and consistency of self-weighing to promote weight-loss maintenance. *Obesity (Silver Spring)* 2020 Jul;28(7):1215-1218 [FREE Full text] [doi: [10.1002/oby.22828](https://doi.org/10.1002/oby.22828)] [Medline: [32437055](https://pubmed.ncbi.nlm.nih.gov/32437055/)]
16. Turicchi J, O'Driscoll R, Finlayson G, Duarte C, Palmeira AL, Larsen SC, et al. Data imputation and body weight variability calculation using linear and nonlinear methods in data collected from digital smart scales: simulation and validation study. *JMIR Mhealth Uhealth* 2020 Sep 11;8(9):e17977 [FREE Full text] [doi: [10.2196/17977](https://doi.org/10.2196/17977)] [Medline: [32915155](https://pubmed.ncbi.nlm.nih.gov/32915155/)]

## Abbreviations

**DEXA:** dual x-ray absorptiometry

**FFI:** foot-to-foot impedance meter

**FFM:** fat-free mass

**FM:** fat mass

**IQR:** interquartile range

*Edited by G Eysenbach; submitted 13.07.20; peer-reviewed by J Rausch, M Amini; comments to author 22.09.20; revised version received 13.11.20; accepted 12.12.20; published 30.04.21.*

*Please cite as:*

*Frija-Masson J, Mullaert J, Vidal-Petiot E, Pons-Kerjean N, Flamant M, d'Ortho MP*

*Accuracy of Smart Scales on Weight and Body Composition: Observational Study*

*JMIR Mhealth Uhealth 2021;9(4):e22487*

*URL: <https://mhealth.jmir.org/2021/4/e22487>*

*doi: [10.2196/22487](https://doi.org/10.2196/22487)*

*PMID: [33929337](https://pubmed.ncbi.nlm.nih.gov/33929337/)*

©Justine Frija-Masson, Jimmy Mullaert, Emmanuelle Vidal-Petiot, Nathalie Pons-Kerjean, Martin Flamant, Marie-Pia d'Ortho. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 30.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Noninvasive Hemoglobin Level Prediction in a Mobile Phone Environment: State of the Art Review and Recommendations

Md Kamrul Hasan<sup>1</sup>, PhD; Md Hasanul Aziz<sup>2</sup>, BSc; Md Ishrak Islam Zarif<sup>2</sup>, BSc; Mahmudul Hasan<sup>3</sup>, BSc; MMA Hashem<sup>4</sup>, PhD; Shion Guha<sup>2</sup>, PhD; Richard R Love<sup>2</sup>, MD; Sheikh Ahamed<sup>2</sup>, PhD

<sup>1</sup>Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, United States

<sup>2</sup>Department of Computer Science, Marquette University, Milwaukee, WI, United States

<sup>3</sup>Department of Computer Science, Stony Brook University, Stony Brook, NY, United States

<sup>4</sup>Department of Computer Science & Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh

**Corresponding Author:**

Md Kamrul Hasan, PhD

Department of Electrical Engineering and Computer Science

Vanderbilt University

334 Featheringill Hall

Nashville, TN

United States

Phone: 1 6153435032

Email: [kamrul.hasan@Vanderbilt.Edu](mailto:kamrul.hasan@Vanderbilt.Edu)

## Abstract

**Background:** There is worldwide demand for an affordable hemoglobin measurement solution, which is a particularly urgent need in developing countries. The smartphone, which is the most penetrated device in both rich and resource-constrained areas, would be a suitable choice to build this solution. Consideration of a smartphone-based hemoglobin measurement tool is compelling because of the possibilities for an affordable, portable, and reliable point-of-care tool by leveraging the camera capacity, computing power, and lighting sources of the smartphone. However, several smartphone-based hemoglobin measurement techniques have encountered significant challenges with respect to data collection methods, sensor selection, signal analysis processes, and machine-learning algorithms. Therefore, a comprehensive analysis of invasive, minimally invasive, and noninvasive methods is required to recommend a hemoglobin measurement process using a smartphone device.

**Objective:** In this study, we analyzed existing invasive, minimally invasive, and noninvasive approaches for blood hemoglobin level measurement with the goal of recommending data collection techniques, signal extraction processes, feature calculation strategies, theoretical foundation, and machine-learning algorithms for developing a noninvasive hemoglobin level estimation point-of-care tool using a smartphone.

**Methods:** We explored research papers related to invasive, minimally invasive, and noninvasive hemoglobin level measurement processes. We investigated the challenges and opportunities of each technique. We compared the variation in data collection sites, biosignal processing techniques, theoretical foundations, photoplethysmogram (PPG) signal and features extraction process, machine-learning algorithms, and prediction models to calculate hemoglobin levels. This analysis was then used to recommend realistic approaches to build a smartphone-based point-of-care tool for hemoglobin measurement in a noninvasive manner.

**Results:** The fingertip area is one of the best data collection sites from the body, followed by the lower eye conjunctival area. Near-infrared (NIR) light-emitting diode (LED) light with wavelengths of 850 nm, 940 nm, and 1070 nm were identified as potential light sources to receive a hemoglobin response from living tissue. PPG signals from fingertip videos, captured under various light sources, can provide critical physiological clues. The features of PPG signals captured under 1070 nm and 850 nm NIR LED are considered to be the best signal combinations following a dual-wavelength theoretical foundation. For error metrics presentation, we recommend the mean absolute percentage error, mean squared error, correlation coefficient, and Bland-Altman plot.

**Conclusions:** We addressed the challenges of developing an affordable, portable, and reliable point-of-care tool for hemoglobin measurement using a smartphone. Leveraging the smartphone's camera capacity, computing power, and lighting sources, we define specific recommendations for practical point-of-care solution development. We further provide recommendations to resolve several long-standing research questions, including how to capture a signal using a smartphone camera, select the best body site

for signal collection, and overcome noise issues in the smartphone-captured signal. We also describe the process of extracting a signal's features after capturing the signal based on fundamental theory. The list of machine-learning algorithms provided will be useful for processing PPG features. These recommendations should be valuable for future investigators seeking to build a reliable and affordable hemoglobin prediction model using a smartphone.

(*JMIR Mhealth Uhealth* 2021;9(4):e16806) doi:[10.2196/16806](https://doi.org/10.2196/16806)

## KEYWORDS

noninvasive hemoglobin; smartphone-based hemoglobin; hemoglobin level from image and video

## Introduction

Hemoglobin (Hb) abnormalities cause several blood diseases, and lead to fatal and chronic health problems, including heart attack, stroke, and pregnancy complications [1]. When an adequate Hb blood level (men  $\geq 13$  g/dL, women  $\geq 12$  g/dL) is not maintained, the disorder complicates the function of the major organs (eg, kidney, brain, and heart) that require oxygen [2]. Anemia, a common Hb disorder, may be caused by blood loss, which is mostly a chronic condition (as occurs with menstruation), decreased red blood cell (RBC) production associated with iron and other nutritional deficiencies, and increased RBC destruction [3,4]. The central role of Hb is to maintain physiologic homeostasis, and the high frequencies of Hb abnormalities make assessment of this parameter a daily clinical activity.

Approximately 5.6% of the US population is anemic and 1.5% of the population has moderate to severe anemia [5]. Sickle cell diseases (SCD) cost more than US \$1.5 billion annually in the United States [6]. Globally, blood disorders and associated complications affect more than 5 million people. In Africa, approximately 250,000 babies are born with SCD every year [7] and 1.62 billion people are affected by Hb-related abnormalities worldwide [8]. A reliable, affordable, and user-friendly solution is crucial to assess the Hb status of a large population. Clinical assessment of Hb typically involves the cyan-methemoglobin method, which is considered to be reliable. However, this invasive process has several limitations, including that the diagnostic devices are not portable, results are not immediately available, and the entire process is expensive. Thus, an Hb disorder diagnosis based on an invasive method is not a perfect solution, especially for people in low- and middle-income countries [9,10]. With available medical facilities, frequent invasive testing is also less convenient due to pain, anxiety, and infections [11]. A recent study estimated the cost for a complete blood count (CBC) test in Bari, Puglia, Italy, with approximately 1,000,000 inhabitants, to be US \$3.14, resulting in a total cost of US \$560,000 in 2018. Considering the entire national territory of Italy, the estimated cost will be more than US \$20 million per year for public hospitals for outpatients. However, the laboratory costs for other cases, including hospitalized patients and private clinic patients, will be much higher than this previous estimation in Italy [12]. These multiple circumstances indicate the reasonable importance of a noninvasive point-of-care (POC) method for Hb measurement.

Commercially available noninvasive POC tools for Hb measurement (Figure 1) are already available [13-16], but have one or more of the following limitations: (1) challenging data collection methods, (2) complex data analysis and feature extraction processes, (3) lack of affordability and portability, and (4) lack of user-friendliness with costly external modules [17]. Smartphone-based solutions are emerging owing to their multifaceted benefits. Recent Hb level assessments use the signal captured from human body locations such as the fingertip [18], nail beds [19], and lower eyelid area [20]. The smartphone's built-in sensors, additional attachments, signal processing methods, and machine-learning algorithms offer major advantages for Hb level estimation. However, most of these components (devices and data collection sites) vary among studies assessing noninvasive methods for Hb level estimation. Therefore, it is important to investigate what, how, and why these components play a vital role in Hb calculation.

Accordingly, in this study, we investigated invasive, minimally invasive, and noninvasive approaches to address the following research questions: (1) How is the signal captured by a smartphone camera from a body site? (2) What issues hinder the smartphone-captured signal for building a noninvasive diagnostic tool? (3) How are a signal's features calculated considering a fundamental theory? (4) What machine-learning algorithms are used to develop a smartphone-based POC diagnostic app?

This study addressed the details of measuring Hb noninvasively. The paper is organized according to the functional components of a noninvasive Hb measurement system. The Methods section describes these components and briefly details current invasive and minimally invasive methodologies for Hb estimation. A list of noninvasive methods is discussed in detail, assessing the challenges and opportunities of smartphone-based solutions. In the Results section, we describe several sensors and signal processing methods that are currently available to process captured signals from different body sites and produce features to apply machine-learning algorithms. We further discuss the most common machine-learning algorithms, including ordinary least squares, multiple linear regression (MLR), partial least square regression (PLSR), and support vector machine regression (SVR). In the Discussion section, we provide several recommendations for the development of a POC tool using a smartphone and propose lighting sources to improve the measurement accuracy levels. Finally, we note our contributions to and limitations of this field.

**Figure 1.** Point-of-care tools for minimally invasive and noninvasive hemoglobin measurement: (a) Hemo Cue, and (b) Astrim-Fit. (These two photos are licensed under CC BY-ND).



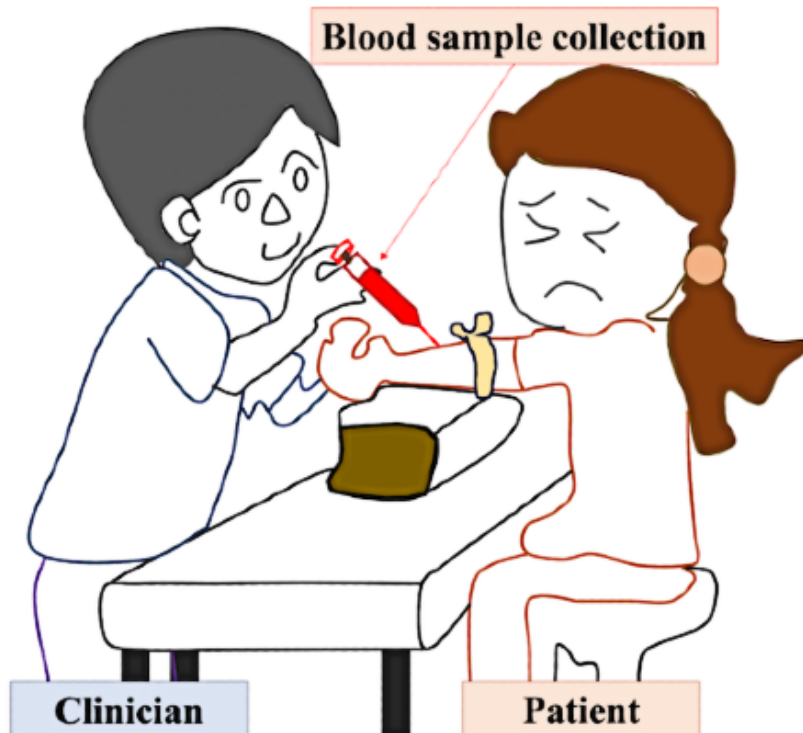
## Methods

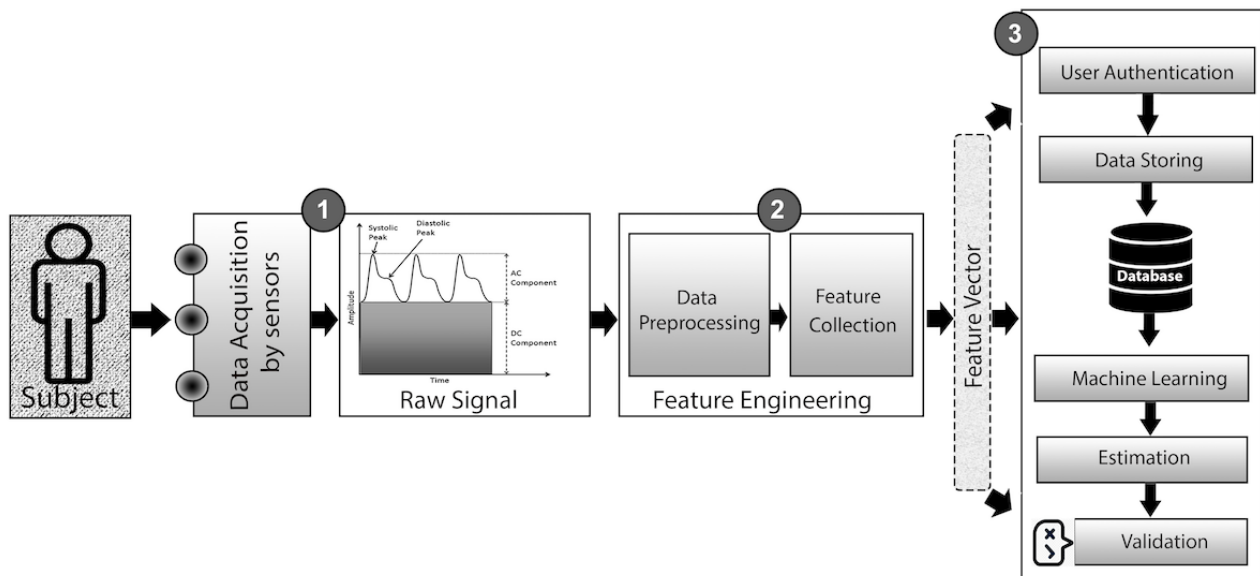
### Overview of Hb Estimation Methods

Hb level measurement is a blood diagnosis process to determine the concentration of Hb in the blood. Clinicians measure Hb in several ways, although the invasive (blood sample collection) approach remains the most common. Invasive processes involve the addition of various chemicals to a blood sample and then optical variations are calculated using spectroscopic data to

measure the Hb level (Figure 2). By contrast, a noninvasive (without blood sample collection) approach involves data obtained from image sensors [21], spectroscopic information, and output of a photoplethysmographic (PPG) sensor to calculate the Hb level (Figure 3). In addition, a minimally invasive process requires only a couple of drops of blood to calculate Hb, and then collects image and spectra-based information from the blood sample for an estimation. Such minimally invasive techniques are comparatively less painful and have fewer complications in collecting sample data.

**Figure 2.** Collecting patient's blood sample for doing invasive hemoglobin diagnosis.



**Figure 3.** Phases involved in a noninvasive hemoglobin measurement system.

### Invasive and Minimally Invasive Processes

Smartphone-based solutions have appeared in recent years for invasive and minimally invasive blood Hb level measurement. In these cases, smartphones follow the characteristics of a spectrometer. For example, Edwards et al [22] built a smartphone-based G-Fresnel spectrometer that works within a wavelength range of 400-1000 nm. The G-Fresnel spectrometer showed a mean error of 9.2% in Hb level measurement in phantom tissue studies. A tungsten halogen lamp illuminates the liquid tissue phantom of human Hb and sends the diffusely reflected light to the G-Fresnel smartphone spectrometer. Although the smartphone-based spectrometer opens a new research horizon toward developing a portable and affordable solution, the use of liquid phantoms may not be an appropriate approach for layered biological tissues. Living tissue has both oxy- and deoxy-Hb molecules, and the phantom's Hb is saturated mostly with oxygen. Therefore, *in vivo* studies seem to be more appropriate for developing a reliable smartphone-based solution to measure Hb levels.

A color-based POC system developed by Tyburski et al [23] was built with an inexpensive, disposable, and standalone device that consisted of two parts: a cap and a body. By capillary action, blood automatically fills the entire sample tube of the body. The cap is then placed into the body, which is prefilled with the reagent solution. After 60 minutes, the blood initiates a redox reaction and the solution shows a stable color change. Using a color scale sticker or with the optional smartphone app after capturing an image of the solution, the Hb level was measured in 238 patients. The sensitivity of the visual interpretation and smartphone analysis of this POC device was 90.2% and 91.1%, respectively, and the specificity was 83.7% and 79.2% respectively. However, this minimally invasive approach suffers from the limitation of the reagent's expiration date, quality of the cap and body of the device, captured image quality and resolution, and identification of an exact Hb level by a visual scale.

Visual image-based approaches have also been introduced for Hb measurement. Glycated Hb (HbA<sub>1c</sub>), which provides information about an average sugar level for the last 4 months, can be measured using a paper-based system and a smartphone, which will help to capture an image of a drop of blood. Using this colorimetric process, Siva et al [24] applied image-processing techniques to investigate the pixel color intensity values and correlated the level of Hb with HbA<sub>1c</sub> [24]. Siemens Healthineers developed and commercially launched a blood diagnosis device named Aina that can be attached to a smartphone for determining levels of Hb, HbA<sub>1c</sub> glucose, and a lipid profile [25].

A chromatography paper-based test was developed that involves a mixture of blood and Drabkin reagent based on bloodstain images digitized with a portable scanner to quantify Hb levels [26]. This process may be accomplished with a smartphone camera sensor by developing a mobile app and analyzing the captured image. For continuous Hb monitoring, a sensor that can calculate Hb levels is implanted in the body, which requires replacement every 3-4 days due to enzyme depletion and membrane contamination [27]. In this implanted system, a wire has to be attached to the patient's body to transmit signals [28].

In these invasive and minimally invasive systems, drawing blood from a vein involves insertion of a needle, which can cause some discomfort, pain, numbness, or a shocking sensation to patients, with subsequent itching or burning at the collection site. These procedures are often traumatic for children and people with mental disabilities. This situation is further exacerbated for patients with needlephobia, a medical condition affecting approximately 10% of the global population [29].

### Noninvasive Process

Noninvasive systems are usually composed of the three main functional components shown in Figure 3: (1) a data acquisition sensor that captures a raw biological (eg, image or spectral) signal; (2) a feature engineering unit that preprocesses the signal and calculates features from the signals; and (3) an Hb level estimation system, which generally incorporates different layers



of user authentication, data storage, prediction model usage, machine learning, and result validation [21]. The user authentication and Hb estimation phase depend on the device, internet availability, user data, and prediction model.

Before elaborating on the smartphone-based noninvasive approaches, we investigated spectroscopy-based techniques with near-infrared (NIR) spectroscopy (NIRS), because these methods have received considerable attention for the noninvasive measurement of blood Hb, oxygenation, pH, hematocrit, and glucose levels [30]. We investigated the lighting sources used in various NIRS methods to determine which NIR lights are most useful in calculating an Hb level noninvasively. The light sources used in a spectra-based investigation may enhance the chance of obtaining accurate Hb information using a smartphone camera to which lights can be attached as an external device.

External lights (ie, NIR lights) are required when a smartphone has no support to sense blood Hb noninvasively in living tissues. One noninvasive method for measuring Hb flow involves analyzing the response of an NIR spectrometer that monitors variations in the absorption of NIR light in the arm, followed by calculating the changes in deoxy-Hb and oxy-Hb concentrations using six wavelengths: 797.5, 802.5, 831.2,

848.7, 866.5, and 907.8 nm [31]. A strong correlation ( $R^2=0.95$ ) between Hb values calculated by venous occlusion PPG and NIRS was determined. Using the PPG signals under eight wavelengths (ranging from 600.22 nm to 1000.60 nm), Yi et al [32] improved the accuracy of dynamic spectrum extraction and analyzed transmitted light through the fingertip of 220 subjects. They developed a calibration model between the dynamic spectrum data and Hb levels, obtaining a correlation of  $r=0.86$  and a root mean square error of prediction of 8.48 g/L. Although the estimated Hb levels were accurate and precise, closely matching clinical requirements, there is an opportunity to involve a more rational calibration set selection process and further improvements of the instrument's signal to noise ratio (SNR). Again, this solution should involve a portable and low-cost instrument.

Table 1 summarizes other spectra-based Hb level measurement processes, which vary in terms of the ranges of light wavelength, input signals, and acquisition devices. In most cases, investigators have used an expensive spectrometer for data collection. Among these spectra-based studies, the most commonly used spectral wavelengths have been 850 nm, 940 nm, and 1070 nm. Investigators have also employed specialized devices to capture PPG signals from the data collection sites such as the finger, hand, and earlobe.

**Table 1.** Summary of spectra-based techniques proposed for noninvasive hemoglobin measurement.

Reference	Wavelength (nm)	Comparator	Signal	Participants (N)
Yi et al [32]	600-1100	Hematology analyzer (Pentra 60; ABX; France)	PPG <sup>a</sup>	220
Rochmanto et al [33]	670, 940	Sysmex-KN21	PPG	78
Desai et al [34]	530	Pronto-7, Hemocue Hb analyzer	PPG	10
Kavsaoglu et al [35]	660, 905	Hemocue Hb-201TM	PPG	33
Kim et al [36]	400-700	Standard CBC test	Photon	32
Nirupa et al [37]	624, 850	Prototype	PPG	69
Ding et al [38]	600-1050	LED <sup>b</sup> and photodiode	Spectra	119
Bremmer et al [39]	350-1050	Ocean Optics DH-2000	Spectra	8
Timm et al [40]	600-1000	LED	PPG	48
Fuksis et al [41]	760-940	IR <sup>c</sup> LEDs	Spectra	— <sup>d</sup>
Pothisarn et al [42]	660, 940	Analyzer oximetry	Light	—
Nguyen et al [43]	940	Radical 7, XE-2100	Pulse	41
Jeon et al [44]	569, 660, 805, 880, 940, 975	Hemoglobin cyanide method	Pulse	129
Jakovels [45]	500-700	White LED	Spectra	—
Timm et al [46]	600-1400	OxyTrue Hb	Spectra	1008
Wang et al [47]	500-700, 1300	Masimo Pronto 7, RGB CMOS camera	PPG	32
Suzaki et al [48]	600, 625, 660, 760, 800, 940, 1300	K1713-09 Hamamatsu Photonics, Co-oximeter	Light	—
Al-Baradie et al [49]	670	Hemo Cue	PPG	10

<sup>a</sup>PPG: photoplethysmography.

<sup>b</sup>LED: light-emitting diode.

<sup>c</sup>IR: infrared.

<sup>d</sup>—: information not provided.

## Results

### Smartphone as a POC Tool

A smartphone-based POC tool as a potential alternative to invasive clinical blood testing is rapidly attracting attention because of the advantages of availability, user-friendliness, and easy attachability to different biosensing devices. The combination of a smartphone and an external device can offer a reliable and affordable POC tool for remote health monitoring. Moreover, the enhanced computing ability, sensing capability, portability, and wide availability of smartphones have propelled this development.

Approximately 56% of US adults and more than 2.5 billion people worldwide are currently using smartphone devices. Multiple critical issues have been addressed in employing smartphones for clinical measurement, including for physiological parameter estimation [50-52], noninvasive Hb diagnosis [53,54], and blood glucose measurement [55,56]. Several studies have shown a higher level of performance in some biomedical applications where a smartphone plays a pivotal role in measuring blood oxygenation, Hb, glucose level, cholesterol, and antibody levels (see [Multimedia Appendix 1](#)). The most frequently used smartphones are developed by Apple, Samsung, Motorola, Google, HTC, Sony, and Asus, where the camera sensor is used to capture videos or images. The accuracy level was deemed to be reliable in each of the studies listed in [Multimedia Appendix 1](#) [50,51,54,55,57-64]. Data commonly captured by a smartphone were obtained from two main body sites: the fingertip and eyelid.

### Finger-Based Analysis

The average width of a human index finger is 14 mm, including the bone (~6 mm), tissue, dermis (~3 mm), epidermis (~1.5 mm), and nail-plate (~1 mm) [65]. As a data collection site, the finger is frequently chosen for several reasons: it is easy to place on a smartphone, it is less sensitive than the eyelid, and it is easy to control. In most cases, the finger pulp area is illuminated using either the phone flashlight or external light sources to obtain the pulsatile information of blood in this area. Reflectance and transmittance oximetry, based on the light source's position, have been applied to the fingertip area using a smartphone to estimate Hb levels. For example, SmartHeLP [53], HemaApp [66], and Hb Meter [67] determinations have used smartphone camera sensors to capture image or videos. In these studies, various lengths of fingertip videos were recorded with different smartphones, and each video frame was analyzed pixel-wise by separating the red, green, and blue (RGB) pixel intensities. Hasan et al [53] subdivided each frame into 10×10 similar sized blocks, separating RGB pixel intensities, and generating time-series information on each block over all frames. They also applied an artificial neural network to estimate Hb levels based on training data of 75 subjects. The gold-standard Hb levels ranged from 7.6 to 13.5 g/dL, and a rank-order correlation of 0.93 was obtained between model-predicted and gold-standard Hb levels. Based on the pixel information from the group of blocks, the most significant region of interest was determined to be close to the smartphone's flashlight. Although RGB pixels were explored in this study, only the red pixel

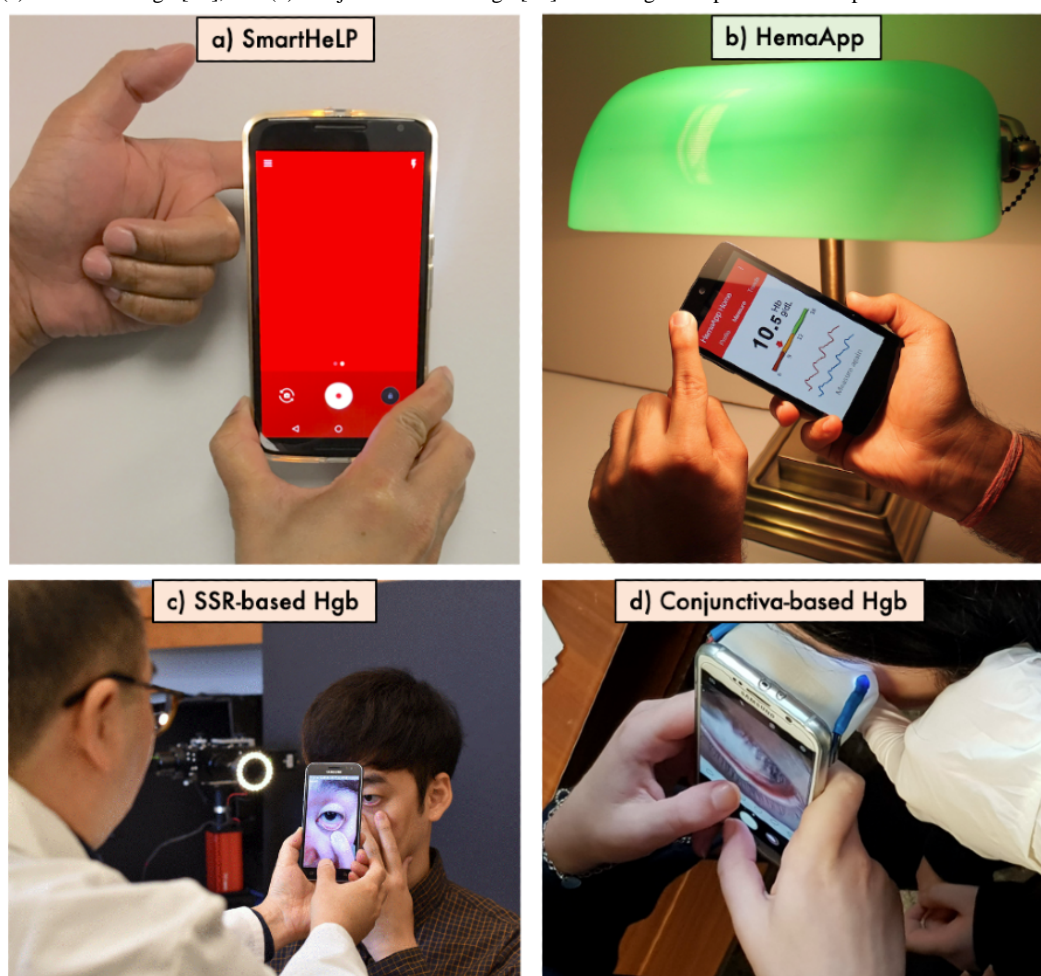
information was employed for development of the prediction model. In addition, the presence of extreme (lower and higher) levels of Hb was limited.

Similarly, Wang et al [66] evaluated fingertip videos using three different hardware embodiments, in which the first embodiment included a white flash and infrared emitter, the second embodiment incorporated an incandescent lamp with a white flash and infrared emitter, and the third embodiment was made by a white flash and custom infrared light-emitting diode (LED) array (Figure 4) [66,68-70]. The external lighting sources, a combination of incandescent and NIR LEDs, resulted in better estimation with an error of 1.26 g/dL and correlation of  $r=0.82$  compared with the other two embodiments. In this case, they captured 15-second-long fingertip videos from 31 subjects, generated pulsatile signals, and extracted RGB time-series waveforms for each video. Additional features, including peak and trough, were calculated from each time-series dataset, and SVR was applied to estimate the level of Hb for each user. In this study, the analyzed Hb levels ranged from 8.3 g/dL to 15.8 g/dL, which were compared with those estimated using Masimo Pronto. Although HemaApp showed greater accuracy than Masimo Pronto, HemaApp was not tested on various types of devices (eg, smartphone) and lighting sources. HemaApp used the bulb to receive light of about 1000 nm, and the age of the light bulb impacted the efficiency. In addition, the effect of ambient light in this study was significant. To make HemaApp more versatile, the prediction model requires upgrading on the training data with more subjects.

In another report, four LED lights with different wavelengths, photodiodes, and a microcontroller unit were used to capture a finger's PPG signal, enabling calculation of the ratio between alternating current (AC) and direct current (DC) signals of the PPG, and estimation of the Hb level, which was transferred to a smartphone through Bluetooth [71]. The microcontroller analyzed the PPG signal using the exponential moving average and then linear regression was applied to calculate the Hb level of 30 subjects, with a root mean square error of 1.53 g/dL. However, the light setting requires correct illumination for precise Hb estimation and the effect of different skin pigmentation is yet to be tested in this system.

A human fingernail, with about 1 mm thickness on average, is comprised of keratin protein, which is translucent [72]. Fingernails have been studied since they allow for easy data capture and they are relatively easy to control [73]. Mannino et al [19] analyzed the images of a fingernail bed captured by a smartphone-based app to investigate critical information for noninvasive Hb level measurement. In this study, an Apple iPhone 5s captured the fingernail bed images (with the camera flash both on and off) from 337 participants who provided blood samples for a standard CBC test. Multilinear regression with a bisquare weighting algorithm was applied to build a prediction model from the nail bed's image parameters and standard laboratory reports. Although the smartphone app measured the Hb level within 2 gm/dL with a bias of 0.2 gm/dL in 100 patients, and showed a good correlation ( $r=0.82$ ) compared with CBC reports, the system suffers from a limitation of automated region of interest selection.

**Figure 4.** Smartphone-based point-of-care tools for noninvasive hemoglobin measurement using fingertip and eyelid images. (a) SmartHeLP [53], b) HemaApp [66], (c) SSR-based Hgb [20], and (d) Conjunctiva-based Hgb [80]. The images are presented with permissions.



Most of the finger-based studies for Hb level estimation considered reflectance oximetry, in which the smartphone camera and the light source were on the same side. However, transmittance oximetry has been rarely applied for finger-based data collection in estimating Hb levels. There is an opportunity to investigate the finger as a data collection site by applying transmittance oximetry, in which light from the finger's dorsal area is sent to the pulp area, because peak absorption of the human melanin pigment occurs at around 335 nm [74], whereas tissue has low absorbance (translucent) in the red and NIR regions, and prior studies indicated that NIR light could penetrate more than 1-2 cm [75].

### Palpebral Conjunctiva

The palpebral conjunctiva, the lower eyelid area of the eye, has received considerable attention as a measurement site because the microvessels in this area are clearly visible and melanocytes are not present [76]. Reflectance spectroscopy has been applied to capture data from the eyelid area [77]. Digital photography and spectral data of lower eyelid images or spectral data converted from an RGB image were applied in several studies to measure Hb levels in a noninvasive manner.

Recently, Park et al [20] introduced a smartphone-based solution that converts an RGB image captured by a smartphone's camera sensor into a virtual hyperspectral image. The need for additional

equipment such as an attachment with a smartphone to capture the spectral response was avoided by generating a conversion matrix ( $T$ ) to transfer a regular image to a spectral image. The generated spectral image was defined as virtual spectra, which were used to train an Hb prediction model. In this study, a wide range of Hb values were calculated in clinical settings and compared with the estimated Hb levels, and Bland-Altman analyses showed reliable performance. Although three different smartphones were used to collect the data, the testing on other smartphones, creation of a more extensive dataset with a wide range of Hb levels, and inclusion of patients with a variety of possible confounding medical conditions are required for further evaluation of this approach.

Selim et al [77] developed a solution based on a lower eyelid image captured by a commercially available Sony DSC-F1 digital camera with a charge-coupled device (CCD), exposing the palpebral conjunctiva. To minimize the effect of ambient light, a gray card was placed close to the eye, and the region of interest was selected from both the eyelid and gray image manually. Invasively measured Hb, using an automated cell counter (SE 9500, Sysmex Corporation, Japan), was compared with the estimated Hb of 117 subjects' eyelid images, demonstrating a Pearson rank-order correlation coefficient of 0.6. However, the prediction algorithm was not verified with other light sources, compared with a gold-standard test, checked

according to the variation in oxygen saturation, or tested in outdoor settings.

Dimauro et al [78] attached an enclosed macrolens with a smartphone to capture a close and high-resolution image of an eyelid with precise focus. The image was segmented using the SLIC Superpixels algorithm, a region of interest was selected for feature extraction, the erythematous was calculated for the CIE-Lab color space, and the k-nearest neighbor classification algorithm [79] was applied to the eyelid image data captured from 102 participants. Applying the Random Oversampling Examples (ROSE) balancing algorithm, they found reliable prediction of the Hb level using conjunctiva images.

Digital images of the palpebral conjunctiva can provide information to measure the level of Hb in a noninvasive manner [80]. Anggraeni et al [57] built a regression model using the digital images of 20 participants' eyelids along with white paper images captured at the same time by Asus ZenFone 2, and estimated the Hb concentration, which correlated highly with clinically measured Hb levels ( $r=0.92$ ). Among the three color pixels of a palpebral conjunctiva image, the red color intensity showed better performance than the green and blue pixel intensities in this study. However, specific software is required for image analysis, and the region of interest needed to be selected for enhancing the precision level.

In addition to conjunctiva images, Rojas et al [81] developed Seliemia, a smartphone and cloud-based platform, using RGB, ISO files, and exposure of images of the tongue, and built a curve-fitting model applying logistic regression and a neural network algorithm. The tongue images provided a better result (sensitivity 91.89% and specificity 85.18%) than the conjunctiva-based prediction model (sensitivity 91.89% and specificity 70.34%) when tested on 64 patients. However, the

training model was built on a population in which most of the participants were young (mean age of 22.6-31.6 years) and extreme levels of Hb were rarely observed in this group (mean 10.6-14.8 g/dL). Establishing a controlled environment and standardized images for Seliemia is challenging.

Although noninvasive devices can capture accurate blood Hb values, their application can be cumbersome and limit users, since the devices have to be attached and oriented correctly, and must be operated with expertise. Access to expensive noninvasive devices for Hb diagnosis is not a practical solution in many low- and middle-income countries. Since the number of smartphone users in the world is estimated at about 6 billion [82], discussion about noninvasive methods should involve user-friendly and cost-effective solutions developed using a smartphone; however, more details of the sensors used are required. In the following sections, we discuss several sensors and signal processing tools.

## Sensors

Sensors translate a physiological signal into machine-accessible data that allow for measurement of physical properties of the human body by collecting physiological signals from one or multiple body sites, including the skin [45], fingertip [66], lip [83], and eye conjunctiva [84]. A machine-learning algorithm with the features generated from a sensor's signal can be used to build a prediction model to estimate Hb levels.

Table 2 lists the different types of sensors that have been used to capture physiological data to estimate Hb levels noninvasively. Most of these sensors are based on image, PPG signal, and optical data. Some of the sensing devices were built by the research team, whereas others used off-the-shelf hardware such as a smartphone, PPG device, or spectrometer.

**Table 2.** Summary of different sensors, signal types, and body sites used for hemoglobin level measurement.

Reference	Device	Sensor	Signal	Body part
Kavsaouglu et al [35]	Hemocue Hb-201	PPG <sup>a</sup>	Light	Finger
Kim et al [36]	Spectrometer, quartz-tungsten-halogen source	Optical	Spectra	Conjunctiva
Nirupa et al [37]	Prototype	PPG	Light	Finger
Ding et al [38]	LED <sup>b</sup> and photodiode	Optical	Spectra	Finger
Timm et al [40]	InGaAs photodiode	Optical	Spectra	Finger
Pothisarn et al [42]	Analyzer oximetry	Optical	Light	Finger
Nguyen et al [43]	XE-2100, Masimo Radical 7	Fluorescence and optical	Pulse	Finger
Jeon et al [44]	Hardware prototype	Optical	Pulse	Finger
Jakovels et al [45]	Nuance 2.4	Optical	Spectra	Skin
Timm et al [46]	Hemocue	Optical	Spectra	Finger
Wang et al [47]	Masimo Pronto 7, RGB <sup>c</sup> CMOS <sup>d</sup> camera	Image	PPG	Fingertip
Kamrul et al [53]	Smartphone camera	Image	PPG	Finger
Wang et al [66]	Smartphone camera	Image	PPG	Finger
Kuestner et al [85]	Modified pulse oximeter, Coulter STKS Monitor	Optical	Spectra	Finger, ear or toe
Lamhaut et al [86]	Hemocue 201+, Radical-7	Optical	Spectra	Finger or ear
Jakovels et al [87]	RGB CMOS	Optical	Spectra	Arm
Miyashita et al [88]	R1-25 and R2-25a	Optical	Spectra	Finger
Li et al [89]	AvaSpec HS1024x58TEC-USB2	Optical	Spectra	Finger
Frasca et al [90]	Hemocue 301, Siemens RapidPoint 405, Sysmex XT 2000i	Optical	Spectra	Finger

<sup>a</sup>PPG: photoplethysmography.

<sup>b</sup>LED: light-emitting diode.

<sup>c</sup>RGB: red, green, blue.

<sup>d</sup>CMOS: complementary metal oxide semiconductor.

Optical sensors, as a type of photometric device, capture the optical signal from an external source such as an LED, laser, or lights of different spectra [91]. Photodiodes are primarily used as optical sensors, which are made of indium gallium arsenide (InGaAs) and indium-phosphor. In some cases, optical sensors contain an embedded amplifier that can select different wavelengths (500-1600 nm) of the signal. A complementary metal oxide semiconductor (CMOS) is a sensor that converts photons to electrons for digital processing, which is used in smartphones, digital video cameras and digital CCTV cameras, astronomical telescopes, scanners, barcode readers, robots, and optical character recognition systems. As high-end smartphone devices include CMOS camera sensors, we can collect data using a smartphone in data collection. In addition to RGB, a CMOS sensor, RGB LED ring-light illuminator, and orthogonally orientated polarizers can be used to capture images, where multiple images under different light sources may carry rare information [87]. The reason is that CMOS chips, with PPG light-capturing cells, pick up the photons at different wavelengths and translate them into electrons, which are converted by digital-to-analog converters into pixels of various

colors [92]. The CCD sensor, a light-sensitive integrated circuit, can convert each image pixel into an electrical charge, and has a high degree of sensitivity that can generate an image even in low-light conditions [93].

InGaAs, an alloy of indium arsenide and gallium arsenide, is another type of infrared sensor used in photodiodes. As a faster response, an InGaAs photodiode is preferred in most studies since these photodiodes have shown higher quantum efficiency [94]. An InGaAs photodetector may also be useful for noninvasively monitoring the Hb concentration and oxygen saturation [95]. A silicon photomultiplier is a solid-state photon detector that can count every single photon, is small in size, of low cost, able to detect low light, and is quantum-efficient [96]. Several studies have used a silicon photomultiplier to build embedded systems to detect PPG signals in both reflective and transmittance modes [97].

A PPG sensor captures an optical response from the microvascular bed of a fingertip, and is used for arterial, venous, and respiratory measurements. Recently, optical PPG sensor data were used for noninvasive Hb measurement [98]. A PPG

device can identify the finger motion with a motion detector and can work with more than one wavelength. Compared with data captured by an electrocardiogram machine, a PPG device reliably ( $R^2=0.93$ ) calculated the heart rate of 170 subjects [99].

Sensing through an electrical-sensing device costs additional money, which complicates the use of such systems. By reducing the number of electrical sensors through leveraging the smartphone's camera sensor, an image or video signal can be collected from a body site, and these data can be processed to generate machine-readable signals and features, and then machine-learning algorithms are applied to build an Hb prediction model. Smartphone-captured data, either image, audio, or video, should be preprocessed using signal processing techniques.

### Signal Processing

Smartphone-recorded signals captured from a body site are attenuated by different unintentional issues such as movement, external noise, and motion artifacts. As part of preprocessing, smartphone-captured data, image, or video are processed using OpenCV library, which generates time-series signals [100] that can be applied to Fourier series analysis on a cycle-by-cycle basis. To remove high-frequency noise in the signal, the data can be filtered using smoothing filters such as Savitzky-Golay smoothing, Butterworth, and Gaussian filters. A cycle-by-cycle Fourier series analysis could reduce the measurement error of the signal from 37% to 3% [101].

The Savitzky-Golay data smoothing filter uses a least-squares polynomial approximation by fitting a polynomial to an input dataset, and evaluates the resulting polynomial at a single point, maintaining the shape and magnitude of the waveform peaks while smoothing the waveform [102].

Biological signals, which are nonstationary as they tend to change over time, can be passed through wavelet transformation for noise reduction and signal enhancement [103]. Stationary wavelet transform was applied to PPG signals, and the wavelet transforms modulus maxima was used to reduce motion artifacts, resulting in an 87% reduction in heart rate estimation error, 76% reduction in heart rate variability estimation error, and 66% reduction in instantaneous heart rate error [104]. A continuous wavelet transform can be used to determine the accurate position of the peak and trough of a PPG signal [105]. However, wave transform has limited capability in restoring corrupted PPG signals for both heart rate and pulse transmit time measurements

[106]. There are also more advanced techniques derived from wave transform such as synchrosqueezing transform that have been used to process PPG signals [107].

Independent component analysis (ICA) can separate the additive non-Gaussian subcomponents of a multivariate signal [108]. As motion artifacts in a PPG signal are derived from independent sources, these can be separated using ICA. ICA can also be used to separate the effect of ambient light and other sources of interference. Kim et al [109] used a combination of ICA and block interleaving with low-pass filtering to reduce motion artifacts in PPG signals. Holton et al [110] compared ICA with principal component analysis, another source separation technique, with respect to their effectiveness in PPG signal recovery from video recordings, and found that ICA produced the most consistent result.

A Butterworth filter, as a maximally flat filter, makes the frequency response of a signal as flat as possible in the passband [111]. By applying the Butterworth filter, high-pass, low-pass, or band-pass filter, a PPG signal can be processed as an authentication method of a PPG biometric [112]. With the Butterworth filter, using both low-pass filtering and wavelet transform, motion artifacts can be removed from PPG data, monitor blood pressure, and identify wrong peaks [113-115].

A biological signal captured by a smartphone introduces noise due to uncontrolled data collection processes, which results in a low SNR. To remove the motion artifact, a Butterworth filter [113], singular value decomposition [116], adaptive filtering [117], Fourier series analysis [118], ICA [109], and principal component analysis [119] have been most commonly used. More than one technique should ideally be used to reduce the motion artifact based on the generated signal's pattern, noise level, sources, environment, and acquisition process. After cleaning, features of the signal are calculated to apply machine-learning algorithms to build a prediction model.

### Machine-Learning Algorithms

#### Definition

A machine-learning algorithm trains a machine to learn and apply acquired knowledge in predictions. Most of the current Hb prediction models use machine-learning algorithms. Although these algorithms could be used in any type of diagnostic system, we here present a list of machine-learning algorithms that are commonly used to assess Hb levels noninvasively (Table 3).

**Table 3.** Summary of machine-learning algorithms for noninvasive hemoglobin measurement.

Reference	Algorithms	Performance measures
Demauro et al [12]	kNN <sup>a</sup> classifier	$r$ and $R^2$
Yi et al [32]	Difference accumulation	$r$ , RMS <sup>b</sup>
Kavsauoglu et al [35]	CART <sup>c</sup> , LSR <sup>d</sup> , GLR <sup>e</sup> , MVLR <sup>f</sup> , PLSR <sup>g</sup> , GRNN <sup>h</sup> , MLR <sup>i</sup> , SVR <sup>j</sup>	MSE <sup>k</sup> , $R^2$ , RMSE <sup>l</sup> , MAPE <sup>m</sup> , IA <sup>n</sup>
Nirupa et al [37]	Linear regression	MSE, $R^2$
Ding et al [38]	BP-ANN <sup>o</sup> and PCA <sup>p</sup>	$r$
Bremmer et al [39]	LLS <sup>q</sup> fit	$r$
Jeon et al [44]	MLR, PLSR	MSE, $R^2$ , $r$
Jakovels et al [45]	Regression analysis	Gaussian analysis
Timm et al [46]	Regression	BAA <sup>r</sup>
Wang et al [47]	Linear regression	RMSE
Wang et al [66]	SVR	$r$ , BAA
Lamhaut et al [86]	Linear regression	$r$ , BAA, $P$ value, bias, precision
Miyashita et al [88]	Linear regression	$r$ and BAA bias plot
Li et al [89]	PLSR	R
Frasca et al [90]	Regression, BAA	MSE, $r$ , RMSE, $R^2$ , BAA
Kamrul et al [18,53]	PLSR and ANN <sup>s</sup>	R

<sup>a</sup>kNN: k-nearest neighbor.

<sup>b</sup>RMS: root mean square.

<sup>c</sup>CART: classification and regression trees.

<sup>d</sup>LSR: least-squares regression.

<sup>e</sup>GLR: generalized linear regression.

<sup>f</sup>MVLR: multivariate linear regression.

<sup>g</sup>PLSR: partial least-squares regression.

<sup>h</sup>GRNN: generalized regression neural network.

<sup>i</sup>MLR: multiple linear regression.

<sup>j</sup>SVR: support vector regression.

<sup>k</sup>MSE: mean square error.

<sup>l</sup>RMSE: root mean square error.

<sup>m</sup>MAPE: mean absolute percentage error.

<sup>n</sup>IA: index of agreement.

<sup>o</sup>BP-ANN: backpropagation artificial neural network.

<sup>p</sup>PCA: principal component analysis.

<sup>q</sup>LLS: linear list squares.

<sup>r</sup>BAA: Bland-Altman analysis.

<sup>s</sup>ANN: artificial neural network.

### MLR

With a similar strategy of more than one simple linear regression, MLR aims to model the relationship between two or more explanatory variables and a response variable. The simple linear regression estimates the relationship between a dependent variable  $Y$  and an explanatory variable  $X$  using the equation  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ , where  $\beta_0$  is the intercept and  $\beta_1$  is the slope of the line, and the error  $\varepsilon_i$  is considered to have a mean value of 0. By contrast, MLR has  $p$  explanatory variables. In

this case, the relationship between  $Y$  and  $X$  is represented by the following equation:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_p X_{pi} + \varepsilon_i,$$

where  $\beta_1$  to  $\beta_p$  are the coefficients. Using this equation, MLR uses the features of a signal as an observation (row) of  $X$  and the target value. For example, the clinically measured Hb level is stored in  $Y$  to build an Hb prediction model [35].

## PLSR

Multiple factors, greater than the number of observations, have been analyzed by PLSR in several studies for Hb level estimation [120,121]. PLSR calculates a few latent factors among other factors that may be responsible for most of the variation in the target or response variable. PLSR names the latent variable as  $T$  or  $X$  scores and defines the response variables as  $U$  or  $Y$  scores. The  $X$  scores with a direction in the factor space explain the factor variation, even when a strong relationship with  $Y$  scores is lacking. In PLSR, the  $Y$  scores maintain the variation of predicted  $Y$  and provide data regarding the change in  $U$ . The greatest advantage of PLSR is that both  $X$  and  $Y$  scores are used to determine a correlation, which helps to build a reliable prediction model [122].

## SVR

SVR is a well-known regression technique for the dimensionality problem, which finds the best hyperplane that separates a class/group with maximum distance using support vectors as a set of critical points. The optimization function is given as follows [123]:

$$1/2 \|w\|^2$$

Subject to,

$$y_i - (w_i x_i) - b < e$$

$$(w_i x_i) + b - y_i < e$$

SVR uses kernels, linear or nonlinear, to create a hyperplane that preserves maximum margins among the data. One of the popular kernel functions is the radial basis function, which has been used to estimate noninvasive Hb levels [35,66,124].

## Measurement Techniques

An Hb prediction model developed by applying a machine-learning algorithm from estimated Hb levels requires a performance test with a gold-standard (clinically measured) Hb value. Performance measurement, based on comparison of estimated with clinically measured values, is generally achieved by calculating the goodness of fit ( $R^2$ ), correlation coefficient ( $r$ ), mean absolute percentage error (MAPE), Bland-Altman plot, mean absolute error, and mean squared error (MSE) in data analysis, as follows.

$$\text{MAPE: } \frac{1}{n} \sum_{i=1}^n \left| \frac{E_i - A_i}{A_i} \right|$$

where  $A_i$  is the actual value or gold-standard measurement,  $E_i$  is the estimated value, and  $n$  is the number of measurements or observations. MAPE is used in the majority of performance measurements because it is easy to explain and understand and does not depend on scale.

If  $Y_i$  denotes the  $i$ th target value and  $\hat{Y}_i$  denotes the estimated value of  $Y_i$ , then the formula for the MSE considering the dependent variable  $Y$  with  $n$  elements is:

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

The correlation coefficient ( $r$ ) demonstrates how strongly two measurement methods are linearly related. The value of  $r$  is

between  $-1.0$  and  $+1.0$ ; if  $r$  is  $+1.0$  or  $-1.0$ , then strong linear relationships are indicated. The formula for Pearson correlation is given by [125]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

where  $n$  is the sample size,  $x_i, y_i$  are the sample points,  $\bar{x}$  is the sample mean, and  $\bar{y}$  is the target mean value.

The Bland-Altman plot is used to estimate a limit of agreement (LOA) between two quantitative measurements. In general, it is common to compute the 95% LOA between two measurement processes. The Bland-Altman plot thus represents the difference between the two measurement methods against the mean value.

## Discussion

### Summary

We investigated several invasive, minimally invasive, and noninvasive methodologies involving a smartphone for data collection, presentation, and transmission processes toward the development of a noninvasive Hb measurement tool. The diverse methodologies across studies included data collection processes, signal processing techniques, feature selection processes, prediction model development algorithms, and performance measurement techniques. Based on these insights, we provide a list of recommendations to develop a smartphone-based noninvasive Hb level estimation tool, which are organized below to answer the research questions on how to capture a signal using a smartphone camera from a body site, address several issues that add noise in the smartphone-captured signal, calculate the features of a signal following a fundamental theory, and apply machine-learning algorithms for the development of an Hb prediction model.

### Body Site Selection for Signal Acquisition

The recommended optimal data collection sites on the body are the palpebral conjunctiva, because of easy access to the microvasculature, and the fingertip, because of the ease of control and access. In the eyelid area, most data collection processes involve digital photography [77,84,126] or reflectance spectroscopy [127,128]. Although most studies demonstrate how to capture an image accurately, perform spectral measurement, and maintain the data collection site motionless during data collection, there is a chance that some of the measurements may include noise from other unintentional activities such as eye blinking, eye sensitivity to the light, breathing, loss of control of the eyelid, or a limited exposed eye area. While using the smartphone camera or external camera to capture an eyelid image, the user can attach a fixed object to the smartphone (eg, mirror) and the image can be captured with a mobile app, in which the boundary of the eyelid area must be visible so that users capture the eyelid image from a fixed distance. In this case, the secondary camera (or the camera on the screen side) of the smartphone is a good option since the user can see the app screen and the eyelid area on the smartphone's screen, which may allow capturing an eyelid image without additional assistance.

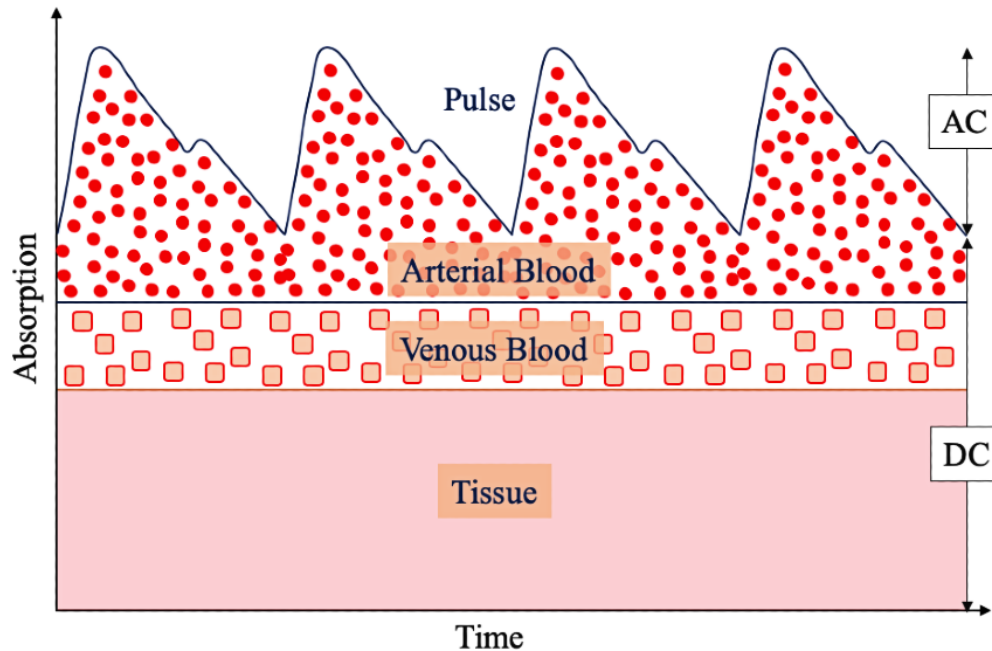


The fingertip has several advantages as a data collection site. Fingertips are easily accessible, less sensitive to minor manipulations, and are generally easy to control. The approximate thickness from the dorsal to the ventral pad side of a finger is about 14 mm for adults. Fingernails with a translucent protein (keratin) can transfer NIR light, which can penetrate more than 1-2 cm [73,75]. Thus, NIR light in the finger tissue can work in both reflective and transmissive mode. Owing to the greater flexibility, we recommend a fingertip-based study over an eyelid-based approach.

## Response Calculation

Fingertip tissues with arterial and venous blood contain light-absorbing components that can be recorded by PPG, an optical device that can be used to observe blood volume changes noninvasively. A PPG system is built with a light source to illuminate the tissue area (eg, finger) and a photodetector to capture the variation of light intensity. The intensity variations are observed due to the systole and diastole parts of the heartbeat. Thus, a PPG signal is derived from two parts: the dynamic part, defined as the AC signal, and the static part, defined as the DC signal (Figure 5).

**Figure 5.** Light absorption changes for pulse, arterial, and venous blood, and living tissue. AC: alternating current; DC: direct current.



The absorption of light by melanin and fat in the skin exhibits a significant response in the shorter wavelengths of light [129]. NIR light, with a wavelength range from 700 to 2500 nm, can penetrate the finger more efficiently than visible light. In this range, light can penetrate 1-2 cm in tissues. As suggested by the foregoing discussion, we recommend light sources of 850 nm or 940 nm, and camera (visible) light to assess the Hb response from a fingertip video. Since water starts showing a response at greater than 950 nm, we recommend using 1070-nm wavelength light to capture the plasma response from a video. The PPG signal captured under an NIR LED light source should be used for further processing. Since the low SNR may reduce the possibility of better PPG generation, selection, and feature extraction, we recommend utilization of laser diodes as a light source and developing a PPG generation algorithm using the dynamic spectrum method [89], ratio of the superimposition averaging template and pulse wave [130], optimized differential extraction method [131], and spectral difference coefficient and dynamic spectrum [132].

## Signal Preprocessing

Because the pressure of the fingertip pad on the smartphone camera and finger movement can alter the waveform of the PPG signal, a well-designed hardware system for securing the imaged finger needs to be developed [133,134]. Noise and artifacts can

be further reduced with the use of filters such as moving average and adaptive filters that work with a reference signal [135]. The reference signals can be obtained from an additional transducer to identify finger movement [136]. Most physiological signals are nonstationary and change their properties over time. In this case, a wavelet transformation and the smoothed pseudo-Wigner-Ville distribution are recommended to improve the PPG signals [135]. The wavelet transform has been used as a common method of movement artifact reduction for PPG signals [137].

To identify the region of interest, HemaApp uses the center section of an image [66], Scully et al [58] used 50×50-sized image pixel intensities on the green channel, and Jonathan and Leahy [138] took a central region with a mean intensity value from 10×10 pixels for smartphone-based PPG generation. Based on these findings, we recommend subdividing an image into a 10×10 image block, generate a PPG signal on each block against all frames, and identify the best location to obtain the strongest PPG signal.

## Theoretical Foundations

The transmissive or reflective process captures the properties of a living tissue noninvasively [139]. The variation of this transmitted or reflected light depends on the shape, volume, and refractive index of Hb, and the angular distribution of

scattered light, which characterizes the absorption properties of blood and tissue [140]. By analyzing these changes in optical scattering properties in tissues, a noninvasive solution for Hb estimation can be achieved.

According to the Beer-Lambert law,  $I_o = I_e e^{-\alpha CD}$ , where  $I_o$  is the output light intensity,  $I$  is the incident light intensity,  $\alpha$  is the light absorption coefficient,  $C$  is the concentration of a blood component, and  $D$  is the light path; the absorption of light is proportional to the concentration of a medium and the path length. A finger has three different absorptions for a given wavelength of light ( $\lambda$ ) due to Hb, plasma ( $P$ ), and the tissue ( $T$ ). Therefore, the light absorption (under a given  $\lambda$ ) by a finger is

$$I_o, \lambda = I_e (\alpha_{Hb}[Hb] + \alpha_P[P] + \alpha_T[T]) (-D)$$

Following the above equation, the light response for the AC and DC value of a PPG can be given as:

$$AC_\lambda = I_e (\alpha_{Hb}[Hb] + \alpha_P[P]) (-d_1) + (\alpha_T[T]) (-DT)$$

$$DC_\lambda = I_e (\alpha_{Hb}[Hb] + \alpha_P[P]) (-d_2) + (\alpha_T[T]) (-DT),$$

where  $d_1$  is the path length for Hb and plasma during the AC signal,  $d_2$  is the path length for Hb and plasma during the DC signal,  $d = d_1 - d_2$ , and  $D_T$  is the path length for the tissue. We assume that the tissue has a stable response, and the ratio of the magnitude of AC and DC removes the effect of the tissue. Then, we can express the ratio between the AC and DC values as:

$$AC_\lambda / DC_\lambda = e^{(\alpha_{Hb}[Hb] + \alpha_P[P]) (-d)}$$

where  $d$  is the path length that affects only the Hb and plasma for  $\lambda$  wavelength of light. Taking the log of both sides of the equation, we can write:

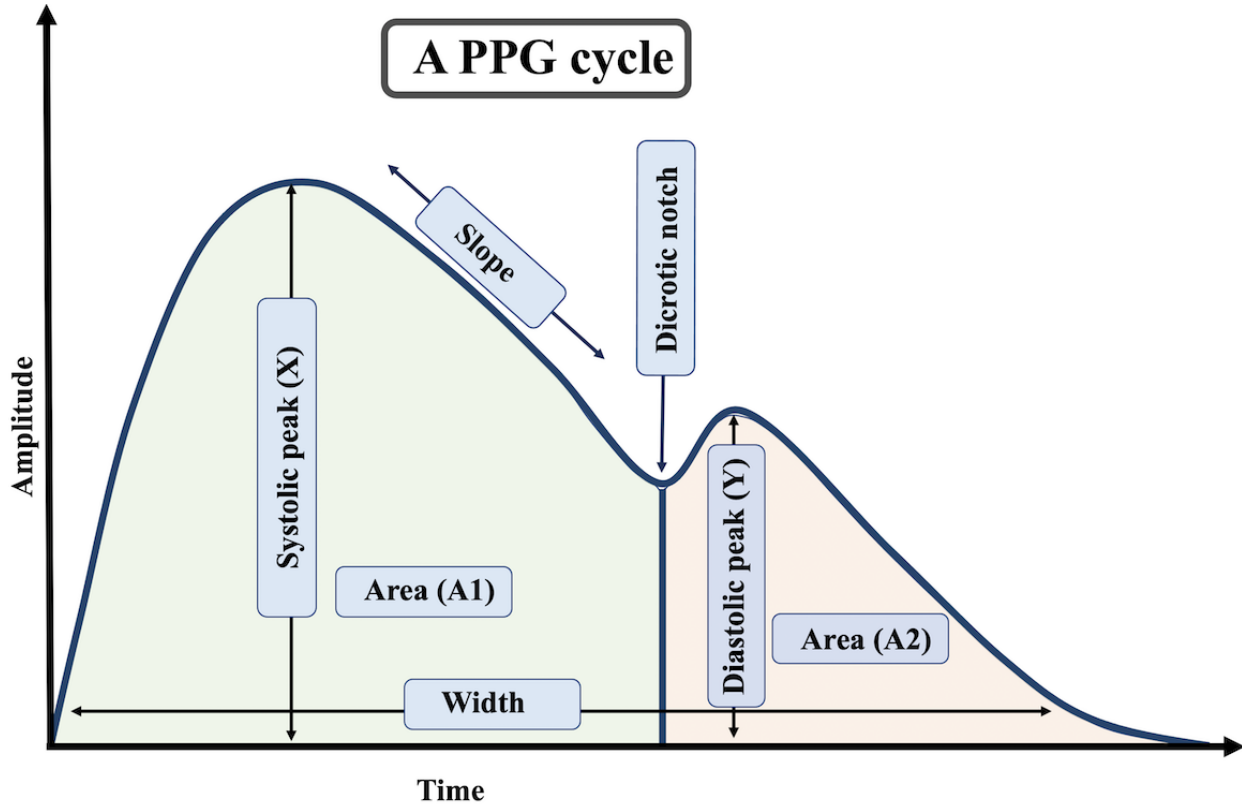
$$\ln AC_\lambda / DC_\lambda = (\alpha_{Hb}[Hb] + \alpha_P[P]) (-d)$$

The empirically measured absorption coefficient for each wavelength of light can help to solve the above equation. However, the system setup for fingertip video recording, lighting conditions, PPG generation from fingertip videos, and complex reflection properties of tissue require machine-learning regression techniques to calculate the ratio of Hb and plasma [54,66]. By incorporating multiple wavelengths of light and the respective responses, several studies have demonstrated reliable Hb prediction models, reducing the number of wavelengths to two with one Hb-sensitive wavelength and another plasma-sensitive wavelength. We recommend this dual-wavelength approach, in which the ratio of the responses captured by two different wavelengths of NIR lights has been applied in different investigations such as for blood Hb [35], skin blood supply assessment [141], oxygenation level [142], and glucose level [143] estimation. The CCD camera sensors can capture PPG signals similar to a pulse oximeter using a photodetector in the NIR range, with a light wavelength around 1000 nm [144]. Modern smartphone cameras have strong sensing capabilities for PPG imaging and volumetric changes in the arterial blood, which enable them to capture PPG signals using reflective or transmissive oximetry from the finger. After generating the PPG signal from the smartphone-based fingertip videos, the features can be calculated from each signal.

### PPG Feature Generation

Since a PPG signal reflects the movement of blood from the heart to the fingertip through the blood vessels, the characteristic parameters of a PPG signal may provide information on blood constituent levels. PPG features have been used in several studies, including those of hematocrit, oxygen saturation, pulse, and respiration [35,145-147]. Based on these insights, we recommend investigating multiple features from the PPG signal (Figure 6), including the systolic and diastolic peak, PPG rise time, pulse transit time, pulse shape, and amplitude [148].

Figure 6. Multiple features collection from a photoplethysmogram signal.



The systolic peak, an indicator of the pulsatile changes in blood volume caused by arterial blood flow, is generated by the direct pressure wave coming from the left ventricle to the periphery of the body. The diastolic peak is a result of reflections of the pressure wave by arteries of the lower body [147]. The dicrotic notch is a small downward deflection between the systolic and diastolic point of a PPG cycle [149]. The pulse interval represents the relationship between the contribution that the wave reflection makes to the systolic arterial pressure and the reflected wave coming from the center [148]. The PPG shows blood movement, whereas the first derivative of the PPG signal indicates the velocity of blood in the finger [150]. Finally, the ratio of a peak value and the sample rate is denoted as the peak time.

The systolic amplitude, representing pulsatile changes in blood volume, can lead a machine-learning algorithm to correlate the pulsatile changes with blood constituent levels [151]. Delle et al [152] confirmed the inverse relationship between the middle cerebral artery peak systolic velocity and Hb levels. With the incoming arterial pulse in the systolic phase, the total light absorbance rises with the increase in arterial blood volume. The systolic increase can then be measured by subtracting the diastolic baseline absorbance from the systolic peak absorbance [153]. The relative augmentation allows us to capture these variations [66], and the inflection points can determine the minimum and maximum values of the PPG waveform [154]. By calculating the first and second derivatives of the PPG signals, the informative inflection points can be more accurately studied. The change in blood volume can be tracked by calculating the pulse interval and the ratio of different peak arrival times [155].

Finally, we recommend calculating the ratio of two PPG features captured under two different wavelengths of light ( $\lambda_H$  and  $\lambda_P$ ). The ratio of two PPG signals' feature values can be presented as follows:

$$R_{\lambda_1}(\lambda_2) = PPG_{\lambda_1} / PPG_{\lambda_2},$$

where,  $R_{\lambda_1}(\lambda_2)$  is the ratio of the two PPG signals' features,  $PPG_{\lambda_1}$  is a PPG generated under the Hb-responsive light source, and  $PPG_{\lambda_2}$  is a PPG calculated under a plasma-responsive light source. The ratio of the two PPG feature values represents the individual ratio between each feature value, which can then be applied to Hb level estimation.

### Dataset Balancing

In medical research, an imbalanced learning problem frequently occurs while solving a classification problem due to insufficient data of certain classes [12]. The imbalance condition can affect the prediction model. Therefore, suitable solutions are required to solve this problem. One strategy might be to alter the class distribution through data resampling (eg, oversampling with sample replacement). The newly generated data can remove the overfitting issues and improve the generalization ability. Class balancing can be achieved through the ROSE algorithm [156], which helps to relieve the severity of the effects of an imbalanced distribution of classes. SMOTE [157], which is based on an oversampling approach, can also be applied to solve this issue.

### Patient Evaluation Strategy

There are different types of users or patients worldwide of a smartphone-based POC solution for blood component measurement. Based on the availability of smartphones and the

expertise of the user, two strategies can be adopted. The first strategy is for users living in low-resource settings, who can obtain the smartphone from a local clinic, pharmacy, village shop, or government office such as a municipality. Since the users are not experts in using the mobile app and face challenges in understanding the output of the blood report, a trained person can help the patient collect the fingertip video or capture an eyelid image to transfer to a cloud for further processing. In this case, the smartphone is safe to use without the risk of losing a device, sending wrong data, and obtaining misleading feedback from the cloud. This option is also cost-effective since many people can access the smartphone with minimum payment. The second strategy is for smartphone users who have some degree of mobile health literacy, confidence to capture data, and a better understanding of mobile apps. In these settings, the users capture data from fingertip videos or eyelid images by themselves and submit the data through the internet. In both contexts, users are also allowed to send their clinical blood test results through the mobile app to the cloud. These strategies will help researchers to build an updated prediction model based on the data stored on the cloud.

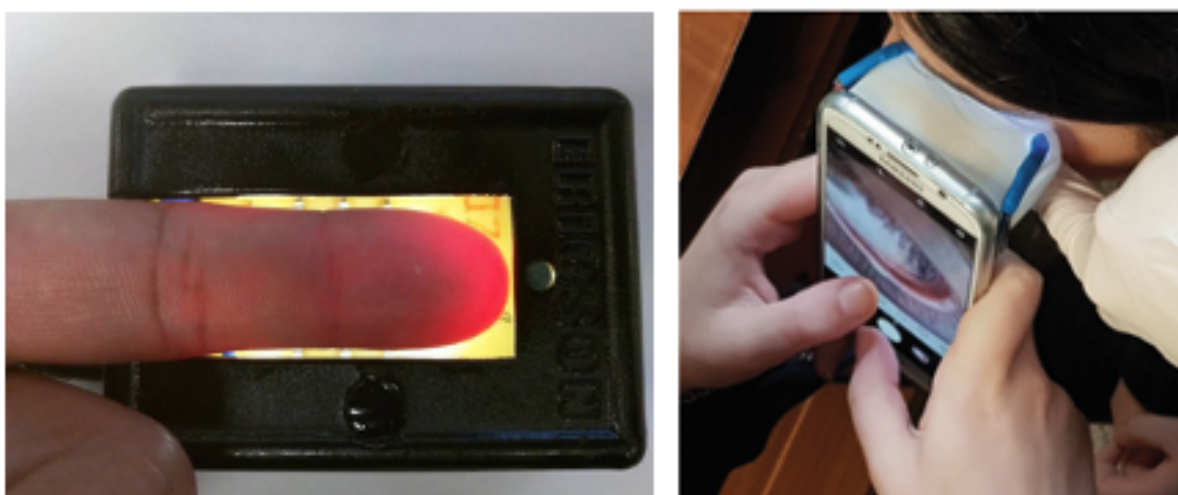
These recommendations can provide guidance for researchers in the area of noninvasive blood component measurement to develop smartphone-based POC tools with the support of mobile app development (user interface), cloud computers, and software and prediction model developers. The data collected by a smartphone can be transferred to a cloud via the internet where several steps are to be accomplished, such as authentication, schedule data to a job manager [158], apply an existing

prediction model, update the model, and give feedback to the users with an estimated Hb level.

## Conclusions

As an increasingly widely available computing platform, the smartphone offers an alternative, noninvasive POC tool to traditional measurements of blood Hb. We recommend the fingertip as the data collection site for the optimal development of an accurate Hb prediction model due to its easy access, use of three different NIR lighting sources, specific signal processing techniques and feature selection methods, and region of interest selection methods. For fingertip-based data collection, a covered external NIR light source (ie, fully covered PPG device) can provide the best PPG signal from a smartphone video. The video should be captured with minimum presence of ambient light, as demonstrated by Hasan et al [159] (Figure 7, left). In addition, an eyelid conjunctiva image can be captured using a smartphone app installed on a head-mounted plastic passive viewer (Figure 7, right) [12,78]. These two data collection methods can provide practical applications because of their reliability, ease of use, and sustainable cost for a patient. Investigators need to consider the following issues before developing such a smartphone-based POC tool: (1) cost of the smartphone, external device, reagents if needed, training, internet, and cloud implementation; (2) other physiological features of the patient; (3) enabling multiple checks with a minimal cognitive load for the user; (4) storing the user's location, sex, and age in the record; (5) keeping the external device as optional so that a user can run a diagnostic without the device; and (6) creating an external device that is cost-effective, easily attachable, properly fit with the finger, and user-friendly.

**Figure 7.** Recommended data collection tool design for (left) fingertip video capture and (right) an eyelid conjunctiva image.



## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Summary of smartphone-based solutions offered in physiological parameter monitoring processes.

[DOCX File , 245 KB - [mhealth\\_v9i4e16806\\_app1.docx](#) ]

## References

1. Ataga KI, Gordeuk VR, Agodoa I, Colby JA, Gittings K, Allen IE. Low hemoglobin increases risk for cerebrovascular disease, kidney disease, pulmonary vasculopathy, and mortality in sickle cell disease: A systematic literature review and meta-analysis. *PLoS One* 2020 Apr 3;15(4):e0229959 [FREE Full text] [doi: [10.1371/journal.pone.0229959](https://doi.org/10.1371/journal.pone.0229959)] [Medline: [32243480](https://pubmed.ncbi.nlm.nih.gov/32243480/)]
2. van Swelm RPL, Wetzels JFM, Swinkels DW. The multifaceted role of iron in renal health and disease. *Nat Rev Nephrol* 2020 Feb 25;16(2):77-98. [doi: [10.1038/s41581-019-0197-5](https://doi.org/10.1038/s41581-019-0197-5)] [Medline: [31554933](https://pubmed.ncbi.nlm.nih.gov/31554933/)]
3. Mann D, Riddell L, Lim K, Byrne LK, Nowson C, Rigo M, et al. Mobile phone app aimed at improving iron intake and bioavailability in premenopausal women: a qualitative evaluation. *JMIR Mhealth Uhealth* 2015 Sep 28;3(3):e92 [FREE Full text] [doi: [10.2196/mhealth.4300](https://doi.org/10.2196/mhealth.4300)] [Medline: [26416479](https://pubmed.ncbi.nlm.nih.gov/26416479/)]
4. Kulandaivelu Y, Lalloo C, Ward R, Zempsky WT, Kirby-Allen M, Breakey VR, et al. Exploring the needs of adolescents with sickle cell disease to inform a digital self-management and transitional care program: qualitative study. *JMIR Pediatr Parent* 2018 Sep 25;1(2):e11058 [FREE Full text] [doi: [10.2196/11058](https://doi.org/10.2196/11058)] [Medline: [31518307](https://pubmed.ncbi.nlm.nih.gov/31518307/)]
5. Le CHH. The prevalence of anemia and moderate-severe anemia in the US population (NHANES 2003-2012). *PLoS One* 2016 Nov 15;11(11):e0166635 [FREE Full text] [doi: [10.1371/journal.pone.0166635](https://doi.org/10.1371/journal.pone.0166635)] [Medline: [27846276](https://pubmed.ncbi.nlm.nih.gov/27846276/)]
6. Krishnamurti L, Ross D, Sinha C, Leong T, Bakshi N, Mittal N, et al. Comparative effectiveness of a web-based patient decision aid for therapeutic options for sickle cell disease: randomized controlled trial. *J Med Internet Res* 2019 Dec 04;21(12):e14462 [FREE Full text] [doi: [10.2196/14462](https://doi.org/10.2196/14462)] [Medline: [31799940](https://pubmed.ncbi.nlm.nih.gov/31799940/)]
7. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults With sickle cell disease: systematic review. *J Med Internet Res* 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: [10.2196/10940](https://doi.org/10.2196/10940)] [Medline: [30026178](https://pubmed.ncbi.nlm.nih.gov/30026178/)]
8. Kawaljit KK. Anaemia 'a silent killer' among women in India: Present scenario. *Eur J Zool Res* 2014;3(1):32-36.
9. Parati G, Stergiou GS, Asmar R, Bilò G, de Leeuw P, Imai Y, ESH Working Group on Blood Pressure Monitoring. European Society of Hypertension practice guidelines for home blood pressure monitoring. *J Hum Hypertens* 2010 Dec 3;24(12):779-785. [doi: [10.1038/jhh.2010.54](https://doi.org/10.1038/jhh.2010.54)] [Medline: [20520631](https://pubmed.ncbi.nlm.nih.gov/20520631/)]
10. Platt OS, Brambilla DJ, Rosse WF, Milner PF, Castro O, Steinberg MH, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med* 1994 Jun 09;330(23):1639-1644. [doi: [10.1056/NEJM199406093302303](https://doi.org/10.1056/NEJM199406093302303)] [Medline: [7993409](https://pubmed.ncbi.nlm.nih.gov/7993409/)]
11. Coventry LL, Jacob AM, Davies HT, Stoneman L, Keogh S, Jacob ER. Drawing blood from peripheral intravenous cannula compared with venepuncture: A systematic review and meta-analysis. *J Adv Nurs* 2019 Nov 20;75(11):2313-2339. [doi: [10.1111/jan.14078](https://doi.org/10.1111/jan.14078)] [Medline: [31115075](https://pubmed.ncbi.nlm.nih.gov/31115075/)]
12. Dimauro G, Guarini A, Caivano D, Girardi F, Pasciolla C, Iacobazzi A. Detecting clinical signs of anaemia from digital images of the palpebral conjunctiva. *IEEE Access* 2019;7:113488-113498. [doi: [10.1109/access.2019.2932274](https://doi.org/10.1109/access.2019.2932274)]
13. Smart HB Hemoglobin Monitor. Enterprises Tech4Life. URL: <https://tech4lifeenterprises.com/smart-hb/> [accessed 2021-03-14]
14. Total Hemoglobin (SpHb). Masimo. URL: <https://www.masimo.com/technology/co-oximetry/sphb/> [accessed 2021-03-14]
15. Kim MJ, Park Q, Kim MH, Shin JW, Kim HO. Comparison of the accuracy of noninvasive hemoglobin sensor (NBM-200) and portable hemoglobinometer (HemoCue) with an automated hematology analyzer (LH500) in blood donor screening. *Ann Lab Med* 2013 Jul 01;33(4):261-267 [FREE Full text] [doi: [10.3343/alm.2013.33.4.261](https://doi.org/10.3343/alm.2013.33.4.261)] [Medline: [23826562](https://pubmed.ncbi.nlm.nih.gov/23826562/)]
16. Sysmex Launches New ASTRIM FIT Product for Noninvasive Measurement of Estimated Hemoglobin Levels. Sysmex. 2014 Jan 15. URL: <https://www.sysmex.co.jp/en/news/2014/140115.html> [accessed 2021-03-13]
17. Bui N, Nguyen A, Nguyen P, Truong H, Ashok A, Dinh T, et al. PhO2: smartphone based blood oxygen level measurement systems using near-IR and RED waveguided light. New York, NY: Association for Computing Machinery; 2017 Presented at: 15th ACM Conference on Embedded Network Sensor Systems; November 6-8, 2017; Delft, Netherlands p. 1-14. [doi: [10.1145/3131672.3131696](https://doi.org/10.1145/3131672.3131696)]
18. Hasan MK, Haque M, Sakib N, Love R, Ahamed SI. Smartphone-based human hemoglobin level measurement analyzing pixel intensity of a fingertip video on different color spaces. *Smart Health* 2018 Jan;5-6:26-39. [doi: [10.1016/j.smhl.2017.11.003](https://doi.org/10.1016/j.smhl.2017.11.003)]
19. Mannino RG, Myers DR, Tyburski EA, Caruso C, Boudreaux J, Leong T, et al. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. *Nat Commun* 2018 Dec 04;9(1):4924. [doi: [10.1038/s41467-018-07262-2](https://doi.org/10.1038/s41467-018-07262-2)] [Medline: [30514831](https://pubmed.ncbi.nlm.nih.gov/30514831/)]
20. Park SM, Visbal-Onufrak MA, Haque MM, Were MC, Naanyu V, Hasan MK, et al. mHealth spectroscopy of blood hemoglobin with spectral super-resolution. *Optica* 2020 Jun 20;7(6):563-573 [FREE Full text] [doi: [10.1364/optica.390409](https://doi.org/10.1364/optica.390409)] [Medline: [33365364](https://pubmed.ncbi.nlm.nih.gov/33365364/)]
21. Blasco J, Chen TM, Tapiador J, Peris-Lopez P. A Survey of Wearable Biometric Recognition Systems. *ACM Comput Surv* 2016 Dec 13;49(3):1-35. [doi: [10.1145/2968215](https://doi.org/10.1145/2968215)]

22. Edwards P, Zhang C, Zhang B, Hong X, Nagarajan VK, Yu B, et al. Smartphone based optical spectrometer for diffusive reflectance spectroscopic measurement of hemoglobin. *Sci Rep* 2017 Sep 22;7(1):12224. [doi: [10.1038/s41598-017-12482-5](https://doi.org/10.1038/s41598-017-12482-5)] [Medline: [28939898](https://pubmed.ncbi.nlm.nih.gov/28939898/)]
23. Tyburski EA, Gillespie SE, Stoy WA, Mannino RG, Weiss AJ, Siu AF, et al. Disposable platform provides visual and color-based point-of-care anemia self-testing. *J Clin Invest* 2014 Oct;124(10):4387-4394. [doi: [10.1172/JCI76666](https://doi.org/10.1172/JCI76666)] [Medline: [25157824](https://pubmed.ncbi.nlm.nih.gov/25157824/)]
24. Vanjari SRK. Minimally invasive mobile healthcare diagnostic platform for glycated hemoglobin detection. *Impacting Research Innovation and Technology (IMPRINT) India*. 2017 Jan 20. URL: <https://imprint-india.org/knowledge-portal-5699-minimally-invasive-mobile-healthcare-diagnostic-platform-for-glycated-hemoglobin-detection> [accessed 2021-03-13]
25. Siemens Launches Mobile Diagnostic Device in India. *The India Expert*. URL: <https://www.theindiaexpert.com/siemens-launches-mobile-diagnostic-device-in-india/> [accessed 2021-03-13]
26. Yang X, Piety NZ, Vignes SM, Benton MS, Kanter J, Shevkopyas SS. Simple paper-based test for measuring blood hemoglobin concentration in resource-limited settings. *Clin Chem* 2013 Oct;59(10):1506-1513 [FREE Full text] [doi: [10.1373/clinchem.2013.204701](https://doi.org/10.1373/clinchem.2013.204701)] [Medline: [23788584](https://pubmed.ncbi.nlm.nih.gov/23788584/)]
27. Sun X, Joseph JI, Crothall KD. Implantable sensor and system for measurement and control of blood constituent levels; US Patent 6,122,536. *Patentscope*. 2000. URL: <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO1997001986> [accessed 2021-03-18]
28. Holz C, Wang EJ. Glabella: Continuously sensing blood pressure behavior using an unobtrusive wearable device. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017 Sep 11;1(3):1-23. [doi: [10.1145/3132024](https://doi.org/10.1145/3132024)]
29. Hamilton J, Ellinwood E. Needle phobia: a neglected diagnosis. *South Med J* 1991;84(Supplement):27. [doi: [10.1097/00007611-199109001-00090](https://doi.org/10.1097/00007611-199109001-00090)]
30. Ellis L, Murphy GJ, Culliford L, Dreyer L, Clayton G, Downes R, et al. The effect of patient-specific cerebral oxygenation monitoring on postoperative cognitive function: a multicenter randomized controlled trial. *JMIR Res Protoc* 2015 Dec 18;4(4):e137 [FREE Full text] [doi: [10.2196/resprot.4562](https://doi.org/10.2196/resprot.4562)] [Medline: [26685289](https://pubmed.ncbi.nlm.nih.gov/26685289/)]
31. Edwards AD, Richardson C, van der Zee P, Elwell C, Wyatt JS, Cope M, et al. Measurement of hemoglobin flow and blood flow by near-infrared spectroscopy. *J Appl Physiol* (1985) 1993 Oct 01;75(4):1884-1889. [doi: [10.1152/jappl.1993.75.4.1884](https://doi.org/10.1152/jappl.1993.75.4.1884)] [Medline: [8282646](https://pubmed.ncbi.nlm.nih.gov/8282646/)]
32. Yi X, Li G, Lin L. Noninvasive hemoglobin measurement using dynamic spectrum. *Rev Sci Instrum* 2017 Aug;88(8):083109. [doi: [10.1063/1.4998978](https://doi.org/10.1063/1.4998978)] [Medline: [28863662](https://pubmed.ncbi.nlm.nih.gov/28863662/)]
33. Rochmanto RA, Zakaria H, Alviana RD, Shahib N. Non-invasive hemoglobin measurement for anemia diagnosis. 2017 Presented at: 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI); September 19-21, 2017; Yogyakarta, Indonesia p. 1-5. [doi: [10.1109/eecsi.2017.8239096](https://doi.org/10.1109/eecsi.2017.8239096)]
34. Desai B, Chaskar U. Comparison of optical sensors for non-invasive hemoglobin measurement. 2016 Presented at: International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT); March 3-5, 2016; Chennai, India p. 2211-2214. [doi: [10.1109/iceeot.2016.7755085](https://doi.org/10.1109/iceeot.2016.7755085)]
35. Kavsaoglu AR, Polat K, Hariharan M. Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features. *Appl Soft Comput* 2015 Dec;37:983-991. [doi: [10.1016/j.asoc.2015.04.008](https://doi.org/10.1016/j.asoc.2015.04.008)]
36. Kim O, McMurdy J, Jay G, Lines C, Crawford G, Alber M. Combined reflectance spectroscopy and stochastic modeling approach for noninvasive hemoglobin determination via palpebral conjunctiva. *Physiol Rep* 2014 Jan 01;2(1):e00192. [doi: [10.1002/phy2.192](https://doi.org/10.1002/phy2.192)] [Medline: [24744871](https://pubmed.ncbi.nlm.nih.gov/24744871/)]
37. Nirupa JLA, Kumar VJ. Non-invasive measurement of hemoglobin content in blood. 2014 Presented at: IEEE International Symposium on Medical Measurements and Applications (MeMeA); June 11-12, 2014; Lisbon, Portugal p. 1-5. [doi: [10.1109/memea.2014.6860140](https://doi.org/10.1109/memea.2014.6860140)]
38. Ding H, Lu Q, Gao H, Peng Z. Non-invasive prediction of hemoglobin levels by principal component and back propagation artificial neural network. *Biomed Opt Express* 2014 Mar 12;5(4):1145-1152. [doi: [10.1364/boe.5.001145](https://doi.org/10.1364/boe.5.001145)]
39. Bremmer RH, Nadort A, van Leeuwen TG, van Gemert MJ, Aalders MC. Age estimation of blood stains by hemoglobin derivative determination using reflectance spectroscopy. *Forensic Sci Int* 2011 Mar 20;206(1-3):166-171. [doi: [10.1016/j.forsciint.2010.07.034](https://doi.org/10.1016/j.forsciint.2010.07.034)] [Medline: [20729018](https://pubmed.ncbi.nlm.nih.gov/20729018/)]
40. Timm U, Andruschenko S, Hinz M. Optical sensor system for continuous non-invasive hemodynamic monitoring in real-time. 2011 Presented at: IEEE Sensors Applications Symposium; February 22-24, 2011; San Antonio, TX p. 167-172. [doi: [10.1109/sas.2011.5739781](https://doi.org/10.1109/sas.2011.5739781)]
41. Fuksis R, Greitans M, Nikisins O. Infrared imaging system for analysis of blood vessel structure. *Elektron Elektrotech* 2010;97(1):45-48.
42. Pothisarn W, Chewpraditkul W, Yupapin PP. Noninvasive hemoglobin-measurement-based pulse oximetry. In: *Optics in Health Care and Biomedical Optics: Diagnostics and Treatment Proceedings Volume 4916*. 2002 Presented at: Photonics Asia; September 12, 2002; Shanghai, China p. 498-504. [doi: [10.1117/12.483014](https://doi.org/10.1117/12.483014)]

43. Nguyen B, Vincent J, Nowak E, Coat M, Paleiron N, Gouny P, et al. The accuracy of noninvasive hemoglobin measurement by multiwavelength pulse oximetry after cardiac surgery. *Anesth Analg* 2011 Nov;113(5):1052-1057. [doi: [10.1213/ANE.0b013e31822c9679](https://doi.org/10.1213/ANE.0b013e31822c9679)] [Medline: [21918163](https://pubmed.ncbi.nlm.nih.gov/21918163/)]
44. Jeon KJ, Kim S, Park KK, Kim J, Yoon G. Noninvasive total hemoglobin measurement. *J Biomed Opt* 2002 Jan;7(1):45-50. [doi: [10.1117/1.1427047](https://doi.org/10.1117/1.1427047)] [Medline: [11818011](https://pubmed.ncbi.nlm.nih.gov/11818011/)]
45. Jakovels D, Spigulis J, Saknite I. Multi-spectral mapping of in vivo skin hemoglobin and melanin. : International Society for Optics and Photonics; 2010 Presented at: Biophotonics: Photonic Solutions for Better Health Care II Volume 7715; April 12-16, 2010; Brussels, Belgium. [doi: [10.1117/12.853928](https://doi.org/10.1117/12.853928)]
46. Timm U, Gewiss H, Kraitl J. Novel multi wavelength sensor concept to detect total hemoglobin concentration, methemoglobin and oxygen saturation. : International Society for Optics and Photonics; 2015 Presented at: Optical Diagnostics and Sensing XV: Toward Point-of-Care Diagnostics Proceedings Volume 9332; March 9, 2015; San Francisco, CA. [doi: [10.1117/12.2080144](https://doi.org/10.1117/12.2080144)]
47. Wang EJ, Li W, Zhu J, Rana R, Patel SN. Noninvasive hemoglobin measurement using unmodified smartphone camera and white flash in Engineering in Medicine and Biology Society (EMBC). : IEEE; 2017 Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 11-15, 2017; Jeju, South Korea p. 2333-2336. [doi: [10.1109/embc.2017.8037323](https://doi.org/10.1109/embc.2017.8037323)]
48. Suzuki H, Kobayashi N, Nagaoka T. Noninvasive measurement of total hemoglobin and hemoglobin derivatives using multiwavelength pulse spectrophotometry-In vitro study with a mock circulatory system. 2006 Presented at: 2006 International Conference of the IEEE Engineering in Medicine and Biology Society; August 30 to September 3, 2006; New York, NY p. 799-802. [doi: [10.1109/iembs.2006.260273](https://doi.org/10.1109/iembs.2006.260273)]
49. Al-Baradie RS, Bose ASC. Portable smart non-invasive hemoglobin measurement system. 2013 Presented at: 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13); March 18-21, 2013; Hammamet, Tunisia p. 1-4. [doi: [10.1109/ssd.2013.6564149](https://doi.org/10.1109/ssd.2013.6564149)]
50. Zaman R, Cho C, Hartmann-Vaccarezza K, Phan T, Yoon G, Chong J. Novel fingertip image-based heart rate detection methods for a smartphone. *Sensors (Basel)* 2017 Feb 12;17(2):358 [FREE Full text] [doi: [10.3390/s17020358](https://doi.org/10.3390/s17020358)] [Medline: [28208678](https://pubmed.ncbi.nlm.nih.gov/28208678/)]
51. Peng R, Zhou X, Lin W, Zhang Y. Extraction of heart rate variability from smartphone photoplethysmograms. *Comput Math Methods Med* 2015;2015:516826. [doi: [10.1155/2015/516826](https://doi.org/10.1155/2015/516826)] [Medline: [25685174](https://pubmed.ncbi.nlm.nih.gov/25685174/)]
52. Kawsar F, Hasan MK, Roushan T, Ahamed SI, Chu WC, Love R. Activity detection using time-delay embedding in multi-modal sensor system. 2016 Presented at: International Conference on Smart Homes and Health Telematics; May 25-27, 2016; Wuhan, China p. 489-499. [doi: [10.1007/978-3-319-39601-9\\_44](https://doi.org/10.1007/978-3-319-39601-9_44)]
53. Hasan MK, Haque MM, Adib R. SmartHeLP: Smartphone-based hemoglobin level prediction using an artificial neural network. : American Medical Informatics Association; 2018 Presented at: AMIA Annual Symposium Proceedings; November 3-7, 2018; San Francisco, USA p. 535.
54. Wang EJ, Zhu J, Li W, Rana R, Patel S. HemaApp IR: noninvasive hemoglobin measurement using unmodified smartphone cameras and built-in LEDs. 2017 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers; Sep 11-15, 2017; Maui Hawaii, USA p. 305-308. [doi: [10.1145/3123024.3123180](https://doi.org/10.1145/3123024.3123180)]
55. Dantu V, Vempati J, Srivilliputhur S. Non-invasive blood glucose monitor based on spectroscopy using a smartphone. 2014 Presented at: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 26-30, 2014; Chicago, IL, USA p. 3695-3698. [doi: [10.1109/embc.2014.6944425](https://doi.org/10.1109/embc.2014.6944425)]
56. Zhang Y, Zhang Y, Siddiqui SA, Kos A. Non-invasive blood-glucose estimation using smartphone PPG signals and subspace kNN classifier. *Elektrotehniski Vestnik* 2019;86(1):68-74.
57. Anggraeni MD, Faton A. Non-invasive self-care anemia detection during pregnancy using a smartphone camera. In: *IOP Conf Ser Mater Sci Eng*. 2017 Feb 17 Presented at: 11th Joint Conference on Chemistry in Conjunction with the 4th Regional Biomaterials Scientific Meeting; September 15-16, 2016; Purwokerto, Indonesia p. 012030. [doi: [10.1088/1757-899x/172/1/012030](https://doi.org/10.1088/1757-899x/172/1/012030)]
58. Scully CG, Lee J, Meyer J, Gorbach AM, Granquist-Fraser D, Mendelson Y, et al. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Trans Biomed Eng* 2012 Feb;59(2):303-306. [doi: [10.1109/tbme.2011.2163157](https://doi.org/10.1109/tbme.2011.2163157)]
59. Gregoski MJ, Mueller M, Vertegel A, Shaporev A, Jackson BB, Frenzel RM, et al. Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. *Int J Telemed Appl* 2012;2012:696324-696237. [doi: [10.1155/2012/696324](https://doi.org/10.1155/2012/696324)] [Medline: [22272197](https://pubmed.ncbi.nlm.nih.gov/22272197/)]
60. Zhu H, Sencan I, Wong J, Dimitrov S, Tseng D, Nagashima K, et al. Cost-effective and rapid blood analysis on a cell-phone. *Lab Chip* 2013 Apr 07;13(7):1282-1288 [FREE Full text] [doi: [10.1039/c3lc41408f](https://doi.org/10.1039/c3lc41408f)] [Medline: [23392286](https://pubmed.ncbi.nlm.nih.gov/23392286/)]
61. Wu Y, Boonloed A, Sleszynski N, Koesdjojo M, Armstrong C, Bracha S, et al. Clinical chemistry measurements with commercially available test slides on a smartphone platform: Colorimetric determination of glucose and urea. *Clin Chim Acta* 2015 Aug 25;448:133-138. [doi: [10.1016/j.cca.2015.05.020](https://doi.org/10.1016/j.cca.2015.05.020)] [Medline: [26102280](https://pubmed.ncbi.nlm.nih.gov/26102280/)]

62. Devadhasan JP, Oh H, Choi CS, Kim S. Whole blood glucose analysis based on smartphone camera module. *J Biomed Opt* 2015 Nov 01;20(11):117001. [doi: [10.1117/1.jbo.20.11.117001](https://doi.org/10.1117/1.jbo.20.11.117001)]
63. Hasan MK, Ahsan GMT, Ahamed SI, Love R, Salim R. Pain level detection from facial image captured by smartphone. *J Inf Process* 2016;24(4):598-608. [doi: [10.2197/ipsjip.24.598](https://doi.org/10.2197/ipsjip.24.598)]
64. Hasan MK, Sakib N, Love RR, Ahamed SI. RGB pixel analysis of fingertip video image captured from sickle cell patient with low and high level of hemoglobin. 2017 Presented at: IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON); October 19-21, 2017; New York City, NY p. 499-505. [doi: [10.1109/uemcon.2017.8249004](https://doi.org/10.1109/uemcon.2017.8249004)]
65. Amaied E, Vargiolu R, Bergheau JM, Zahouani H. Aging effect on tactile perception: Experimental and modelling studies. *Wear* 2015 May;332-333:715-724. [doi: [10.1016/j.wear.2015.02.030](https://doi.org/10.1016/j.wear.2015.02.030)]
66. Wang EJ, Li W, Hawkins D, Gernsheimer T, Norby-Slycord C, Patel SN. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. 2016 Presented at: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12, 2016; Heidelberg, Germany p. 593-604. [doi: [10.1145/2971648.2971653](https://doi.org/10.1145/2971648.2971653)]
67. Flash of insight: mobile app tests for anemia with camera flash. *Africare In the News*. 2015 Jun 04. URL: <https://www.africare.org/flash-of-insight-mobile-app-tests-for-anemia-with-camera-flash/> [accessed 2021-03-14]
68. Why track oxygen content? *Ember*. URL: <https://www.cercacor.com/pages/ember> [accessed 2021-03-14]
69. Effective and affordable treatment for non-life threatening illnesses. *Biosense*. URL: <https://beta.biosense.in/services/diagnostic-center/> [accessed 2021-03-14]
70. Chansanchai T. Imagine Cup World Finals 2014 winner Eyaemia receives \$50,000 and a private meeting with Bill Gates. *Official Microsoft Blog*. 2014 Aug 01. URL: <https://blogs.microsoft.com/blog/2014/08/01/imagine-cup-world-finals-2014-winner-eyenaemia-receives-50000-private-meeting-bill-gates/> [accessed 2021-03-14]
71. Dewantoro P, Gandana CE, Zakaria RORH, Irawan YS. Development of smartphone-based non-invasive hemoglobin measurement. 2018 Presented at: 2018 International Symposium on Electronics and Smart Devices (ISESD); October 23-24, 2018; Bandung, Indonesia p. 1-6. [doi: [10.1109/isesd.2018.8605489](https://doi.org/10.1109/isesd.2018.8605489)]
72. Miki RA. Nail plate growth guide surgical implantation kit US Patent 10,806,492. USPTO. 2020. URL: <https://uspto.report/patent/grant/10,806,492> [accessed 2020-03-18]
73. Hayashi M, Yamamoto K, Yoshimura M, Hayashi H, Shitara A. Cadmium, lead, and zinc concentrations in human fingernails. *Bull Environ Contam Toxicol* 1993 Apr;50(4). [doi: [10.1007/bf00191244](https://doi.org/10.1007/bf00191244)]
74. Kollias N. The spectroscopy of human melanin pigmentation, melanin: its role in human photoprotection. *Valden Pub* 1995:31-38.
75. Carroll L, Humphreys TR. LASER-tissue interactions. *Clin Dermatol* 2006 Jan;24(1):2-7. [doi: [10.1016/j.clindermatol.2005.10.019](https://doi.org/10.1016/j.clindermatol.2005.10.019)] [Medline: [16427500](https://pubmed.ncbi.nlm.nih.gov/16427500/)]
76. Lin JY, Fisher DE. Melanocyte biology and skin pigmentation. *Nature* 2007 Feb 22;445(7130):843-850. [doi: [10.1038/nature05660](https://doi.org/10.1038/nature05660)] [Medline: [17314970](https://pubmed.ncbi.nlm.nih.gov/17314970/)]
77. Suner S, Crawford G, McMurdy J, Jay G. Non-invasive determination of hemoglobin by digital photography of palpebral conjunctiva. *J Emerg Med* 2007 Aug;33(2):105-111. [doi: [10.1016/j.jemermed.2007.02.011](https://doi.org/10.1016/j.jemermed.2007.02.011)] [Medline: [17692757](https://pubmed.ncbi.nlm.nih.gov/17692757/)]
78. Dimauro G, Baldari L, Caivano D, Colucci G, Girardi F. Automatic segmentation of relevant sections of the conjunctiva for non-invasive anemia detection. 2018 Presented at: 3rd International Conference on Smart and Sustainable Technologies (SpliTech); June 26-29, 2018; Split, Croatia p. 1-5.
79. Deng Z, Zhu X, Cheng D, Zong M, Zhang S. Efficient k NN classification algorithm for big data. *Neurocomputing* 2016 Jun;195:143-148. [doi: [10.1016/j.neucom.2015.08.112](https://doi.org/10.1016/j.neucom.2015.08.112)]
80. Bevilacqua V, Dimauro G, Marino F. A novel approach to evaluate blood parameters using computer vision techniques. : IEEE; 2016 Presented at: IEEE International Symposium on Medical Measurements and Applications (MeMeA); May 15-18, 2016; Benevento, Italy p. 1-6. [doi: [10.1109/memea.2016.7533760](https://doi.org/10.1109/memea.2016.7533760)]
81. Wightman Rojas PM, Mass Noriega LA, Salazar Silva A. Hemoglobin screening using cloud based mobile photography applications. *Ingener Universidad* 2019 Jun 21;23(2). [doi: [10.11144/javeriana.iyu23-2.hsuc](https://doi.org/10.11144/javeriana.iyu23-2.hsuc)]
82. Gozalvez J. Samsung Electronics Sets 5G Speed Record at 7.5 Gb/s. *IEEE Veh Technol Mag* 2015 Mar;10(1):12-16. [doi: [10.1109/mvt.2015.2390931](https://doi.org/10.1109/mvt.2015.2390931)]
83. Hennig G, Homann C, Teksan I, Hasbargen U, Hasmüller S, Holdt LM, et al. Non-invasive detection of iron deficiency by fluorescence measurement of erythrocyte zinc protoporphyrin in the lip. *Nat Commun* 2016 Feb 17;7(1):10776. [doi: [10.1038/ncomms10776](https://doi.org/10.1038/ncomms10776)] [Medline: [26883939](https://pubmed.ncbi.nlm.nih.gov/26883939/)]
84. Collings S, Thompson O, Hirst E, Goossens L, George A, Weinkove R. Non-invasive detection of anaemia using digital photographs of the conjunctiva. *PLoS One* 2016 Apr 12;11(4):e0153286 [FREE Full text] [doi: [10.1371/journal.pone.0153286](https://doi.org/10.1371/journal.pone.0153286)] [Medline: [27070544](https://pubmed.ncbi.nlm.nih.gov/27070544/)]
85. Kuestner JT. Method for non-invasive and in-vitro hemoglobin concentration measurement US Patent 5,377,674. *Unified Patents Portal*. 2015. URL: <https://portal.unifiedpatents.com/patents/patent/US-5377674-A> [accessed 2021-03-18]



86. Lamhaut L, Apriotesei R, Combes X, Lejay M, Carli P, Vivien B. Comparison of the accuracy of noninvasive hemoglobin monitoring by spectrophotometry (SpHb) and HemoCue® with automated laboratory hemoglobin measurement. *Anesthesiology* 2011 Sep;115(3):548-554 [FREE Full text] [doi: [10.1097/ALN.0b013e3182270c22](https://doi.org/10.1097/ALN.0b013e3182270c22)] [Medline: [21716091](https://pubmed.ncbi.nlm.nih.gov/21716091/)]
87. Jakovels D, Spigulis J. RGB imaging device for mapping and monitoring of hemoglobin distribution in skin. *Lith J Phys* 2012;52(1):50-54. [doi: [10.3952/physics.v52i1.2267](https://doi.org/10.3952/physics.v52i1.2267)]
88. Miyashita R, Hirata N, Sugino S, Mimura M, Yamakage M. Improved non-invasive total haemoglobin measurements after in-vivo adjustment. *Anaesthesia* 2014 Jul 07;69(7):752-756. [doi: [10.1111/anae.12681](https://doi.org/10.1111/anae.12681)] [Medline: [24800903](https://pubmed.ncbi.nlm.nih.gov/24800903/)]
89. Li G, Xu S, Zhou M, Zhang Q, Lin L. Noninvasive hemoglobin measurement based on optimizing Dynamic Spectrum method. *Spectrosc Lett* 2017 May 05;50(3):164-170. [doi: [10.1080/00387010.2017.1302481](https://doi.org/10.1080/00387010.2017.1302481)]
90. Frasca D, Dahyot-Fizelier C, Catherine K, Levrat Q, Debaene B, Mimoz O. Accuracy of a continuous noninvasive hemoglobin monitor in intensive care unit patients. *Crit Care Med* 2011 Oct;39(10):2277-2282. [doi: [10.1097/CCM.0b013e3182227e2d](https://doi.org/10.1097/CCM.0b013e3182227e2d)] [Medline: [21666449](https://pubmed.ncbi.nlm.nih.gov/21666449/)]
91. Kraithl J, Timm U, Ewald H, Lewis E. Non-invasive sensor for an in vivo hemoglobin measurement. 2011 Presented at: Sensors IEEE; October 28-31, 2011; Limerick, Ireland p. 276-279. [doi: [10.1109/icsens.2011.6126982](https://doi.org/10.1109/icsens.2011.6126982)]
92. CMOS sensor. TechTarget. URL: <https://whatis.techtarget.com/definition/CMOS-sensor> [accessed 2021-03-14]
93. Charge-coupled device (CCD). TechTarget. URL: <https://searchstorage.techtarget.com/definition/charge-coupled-device> [accessed 2021-03-14]
94. Eppeldauer GP. Electronic characteristics of Ge and InGaAs radiometers. In: Proceedings of SPIE 3061, Infrared Technology and Applications XXIII.: SPIE; 1997 Presented at: Proceedings of SPIE 3061, Infrared Technology and Applications XXIII; Aug 13, 1997; Orlando, FL p. 833-838.
95. Jeon K, Yoon G. Method and apparatus for noninvasively monitoring hemoglobin concentration and oxygen saturation US Patent 6,714,805. Google Patents. 2004. URL: <https://patents.google.com/patent/US20030009090> [accessed 2021-03-18]
96. Agrò D, Canicatti R, Tomasino A. PPG embedded system for blood pressure monitoring. : IEEE; 2014 Presented at: AEIT Annual Conference From Research to Industry: The Need for a More Effective Technology Transfer (AEIT 2014); September 18-19, 2014; Trieste, Italy p. 1-6. [doi: [10.1109/aeit.2014.7002012](https://doi.org/10.1109/aeit.2014.7002012)]
97. Rundo F, Conoci S, Ortis A, Battiato S. An Advanced Bio-Inspired PhotoPlethysmoGraphy (PPG) and ECG Pattern Recognition System for Medical Assessment. *Sensors (Basel)* 2018 Jan 30;18(2):405 [FREE Full text] [doi: [10.3390/s18020405](https://doi.org/10.3390/s18020405)] [Medline: [29385774](https://pubmed.ncbi.nlm.nih.gov/29385774/)]
98. Phillips JP, Hickey M, Kyriacou PA. Evaluation of electrical and optical plethysmography sensors for noninvasive monitoring of hemoglobin concentration. *Sensors (Basel)* 2012 Feb 09;12(2):1816-1826 [FREE Full text] [doi: [10.3390/s120201816](https://doi.org/10.3390/s120201816)] [Medline: [22438739](https://pubmed.ncbi.nlm.nih.gov/22438739/)]
99. Uguz DU, Venema B, Leonhardt S, Teichmann D. Multifunctional photoplethysmography sensor design for respiratory and cardiovascular diagnosis. In: World Congress on Medical Physics and Biomedical Engineering 2018. Prague, Czech Republic: Springer; 2018:905-909.
100. Howse J. OpenCV computer vision with python. In: Packt Publishing Ltd. Birmingham, United Kingdom: Packt Publishing Ltd; 2013.
101. Reddy K, George B, Kumar V. Use of Fourier Series Analysis for Motion Artifact Reduction and Data Compression of Photoplethysmographic Signals. *IEEE Trans Instrum Meas* 2009 May;58(5):1706-1711. [doi: [10.1109/tim.2008.2009136](https://doi.org/10.1109/tim.2008.2009136)]
102. Schafer R. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Process Mag* 2011 Jul;28(4):111-117. [doi: [10.1109/msp.2011.941097](https://doi.org/10.1109/msp.2011.941097)]
103. Tamura T, Maeda Y, Sekine M, Yoshida M. Wearable Photoplethysmographic Sensors—Past and Present. *Electronics* 2014 Apr 23;3(2):282-302. [doi: [10.3390/electronics3020282](https://doi.org/10.3390/electronics3020282)]
104. Lee C. Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach. 2003 Presented at: IEEE EMBS Asian-Pacific Conference on Biomedical Engineering; October 20-22, 2003; Kyoto-Osaka-Nara, Japan p. 194-195. [doi: [10.1109/apbme.2003.1302650](https://doi.org/10.1109/apbme.2003.1302650)]
105. Teng X, Zhang Y. Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach. 2003 Presented at: 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439); September 17-21, 2003; Cancun, Mexico p. 3153-3156. [doi: [10.1109/iembs.2003.1280811](https://doi.org/10.1109/iembs.2003.1280811)]
106. Foo JYA. Comparison of wavelet transformation and adaptive filtering in restoring artefact-induced time-related measurement. *Biomed Signal Process Control* 2006 Jan;1(1):93-98. [doi: [10.1016/j.bspc.2006.01.001](https://doi.org/10.1016/j.bspc.2006.01.001)]
107. Dehkordi P, Garde A, Molavi B, Ansermino JM, Dumont GA. Extracting instantaneous respiratory rate from multiple photoplethysmogram respiratory-induced variations. *Front Physiol* 2018 Jul 18;9:948. [doi: [10.3389/fphys.2018.00948](https://doi.org/10.3389/fphys.2018.00948)] [Medline: [30072918](https://pubmed.ncbi.nlm.nih.gov/30072918/)]
108. Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. *Neural Netw* 2000 Jun;13(4-5):411-430. [doi: [10.1016/s0893-6080\(00\)00026-5](https://doi.org/10.1016/s0893-6080(00)00026-5)]
109. Kim B, Yoo S. Motion Artifact Reduction in Photoplethysmography Using Independent Component Analysis. *IEEE Trans Biomed Eng* 2006 Mar;53(3):566-568. [doi: [10.1109/tbme.2005.869784](https://doi.org/10.1109/tbme.2005.869784)]

110. Holton BD, Mannapperuma K, Lesniewski PJ, Thomas JC. Signal recovery in imaging photoplethysmography. *Physiol Meas* 2013 Nov 22;34(11):1499-1511. [doi: [10.1088/0967-3334/34/11/1499](https://doi.org/10.1088/0967-3334/34/11/1499)] [Medline: [24149772](https://pubmed.ncbi.nlm.nih.gov/24149772/)]
111. Zumbahlen H. *Linear circuit design handbook*. Oxford, UK: Newnes Elsevier; 2011.
112. Bonissi A, Labati RD, Perico L, Sassi R, Scotti F, Sparagino L. A preliminary study on continuous authentication methods for photoplethysmographic biometrics. 2013 Presented at: 2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications; Oct 17, 2014; Napoli, Italy p. 28-33. [doi: [10.1109/bioms.2013.6656145](https://doi.org/10.1109/bioms.2013.6656145)]
113. Luke A, Shaji S, Menon U. Performance enhancement of a photoplethysmographic biosensor using efficient signal processing techniques. 2018 Presented at: 3rd International Conference for Convergence in Technology (I2CT); April 6-8, 2018; Pune, India p. 1-6. [doi: [10.1109/i2ct.2018.8529566](https://doi.org/10.1109/i2ct.2018.8529566)]
114. Riaz F, Azad MA, Arshad J, Imran M, Hassan A, Rehman S. Pervasive blood pressure monitoring using Photoplethysmogram (PPG) sensor. *Future Gener Comput Syst* 2019 Sep;98:120-130. [doi: [10.1016/j.future.2019.02.032](https://doi.org/10.1016/j.future.2019.02.032)]
115. Moreno S, Quintero-Parra A, Ochoa-Pertuz C, Villarreal R, Kuzmar I. A signal processing method for respiratory rate estimation through photoplethysmography. *Int J Sign Process Image Process Patt Recogn* 2018 Apr 30;11(2):1-10. [doi: [10.14257/ijsip.2018.11.2.01](https://doi.org/10.14257/ijsip.2018.11.2.01)]
116. Reddy KA, Kumar VJ. Motion artifact reduction in photoplethysmographic signals using singular value decomposition. 2007 Presented at: IEEE Instrumentation & Measurement Technology Conference IMTC; May 1-3, 2007; Warsaw, Poland p. 1-4. [doi: [10.1109/imtc.2007.379467](https://doi.org/10.1109/imtc.2007.379467)]
117. Ram MR, Madhav KV, Krishna EH, Komalla NR, Reddy KA. A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Trans Instrum Meas* 2012 May;61(5):1445-1457. [doi: [10.1109/tim.2011.2175832](https://doi.org/10.1109/tim.2011.2175832)]
118. Reddy K, George B, Kumar V. Use of Fourier series analysis for motion artifact reduction and data compression of photoplethysmographic signals. *IEEE Trans Instrum Meas* 2009 May;58(5):1706-1711. [doi: [10.1109/tim.2008.2009136](https://doi.org/10.1109/tim.2008.2009136)]
119. Ram MR, Madhav KV, Krishna EH, Reddy KN, Reddy KA. Use of multi-scale principal component analysis for motion artifact reduction of PPG signals. 2011 Presented at: IEEE Recent Advances in Intelligent Computational Systems; September 22, 2011; Trivandrum, India p. 425-430. [doi: [10.1109/raics.2011.6069348](https://doi.org/10.1109/raics.2011.6069348)]
120. Mohd Idrus MNE, Chia KS. Partial least square with Savitzky Golay derivative in predicting blood hemoglobin using near infrared spectrum. *MATEC Web of Conferences* 2018 Feb 23;150:01001. [doi: [10.1051/mateconf/201815001001](https://doi.org/10.1051/mateconf/201815001001)]
121. Pan T, Yan B, Tang Y, Chen J, Yao L. Combination optimization of PLS regression and SG smoothing in NIR analysis of hemoglobin. 2017 Presented at: International Conference on Photonics and Imaging in Biology and Medicine; September 26-28, 2017; Suzhou, China. [doi: [10.1364/pibm.2017.w3a.75](https://doi.org/10.1364/pibm.2017.w3a.75)]
122. Tobias RD. *An introduction to partial least squares regression*. : SAS Institute Inc Cary; 1995 Presented at: Proceedings of the twentieth annual SAS users group international conference; 1995; Orlando, Florida.
123. Awad M, Khanna R. Support vector regression. In: *Efficient learning machines*. Berkeley, CA: Apress; 2015:67-80.
124. Naseer N, Hong K, Bhutta MR, Khan MJ. Improving classification accuracy of covert yes/no response decoding using support vector machines: an fNIRS study. 2014 Presented at: International Conference on Robotics and Emerging Allied Technologies in Engineering (iCREATE); April 22-24, 2014; Islamabad, Pakistan p. 6-9. [doi: [10.1109/icreate.2014.6828329](https://doi.org/10.1109/icreate.2014.6828329)]
125. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012 Sep;24(3):69-71 [FREE Full text] [Medline: [23638278](https://pubmed.ncbi.nlm.nih.gov/23638278/)]
126. Kim T, Choi SH, Lambert-Cheatham N, Xu Z, Kritchevsky JE, Bertin F, et al. Toward laboratory blood test-comparable photometric assessments for anemia in veterinary hematology. *J Biomed Opt* 2016 Oct 01;21(10):107001. [doi: [10.1117/1.JBO.21.10.107001](https://doi.org/10.1117/1.JBO.21.10.107001)] [Medline: [27704141](https://pubmed.ncbi.nlm.nih.gov/27704141/)]
127. Sarkar PK, Pal S, Polley N, Aich R, Adhikari A, Halder A, et al. Development and validation of a noncontact spectroscopic device for hemoglobin estimation at point-of-care. *J Biomed Opt* 2017 May 01;22(5):55006. [doi: [10.1117/1.JBO.22.5.055006](https://doi.org/10.1117/1.JBO.22.5.055006)] [Medline: [28510622](https://pubmed.ncbi.nlm.nih.gov/28510622/)]
128. Khansari MM, Tan M, Karamian P, Shahidi M. Inter-visit variability of conjunctival microvascular hemodynamic measurements in healthy and diabetic retinopathy subjects. *Microvasc Res* 2018 Jul;118:7-11 [FREE Full text] [doi: [10.1016/j.mvr.2018.01.007](https://doi.org/10.1016/j.mvr.2018.01.007)] [Medline: [29438814](https://pubmed.ncbi.nlm.nih.gov/29438814/)]
129. Jones DP. Medical electro-optics: measurements in the human microcirculation. *Physics Technol* 2002 Nov 20;18(2):79-85. [doi: [10.1088/0305-4624/18/2/305](https://doi.org/10.1088/0305-4624/18/2/305)]
130. Li G, Xiong C, Wang H, Lin L, Zhang B, Tong Y. Single-trial estimation of dynamic spectrum. *Guang Pu Xue Yu Guang Pu Fen Xi* 2011 Jul;31(7):1857-1861. [Medline: [21942039](https://pubmed.ncbi.nlm.nih.gov/21942039/)]
131. Tang W, Chen Q, Yan W, He G, Li G, Lin L. An Optimizing Dynamic Spectrum Differential Extraction Method for Noninvasive Blood Component Analysis. *Appl Spectrosc* 2020 Jan 12;74(1):23-33. [doi: [10.1177/0003702818815508](https://doi.org/10.1177/0003702818815508)] [Medline: [30409032](https://pubmed.ncbi.nlm.nih.gov/30409032/)]
132. Feng X, Li G, Yu H, Wang S, Yi X, Lin L. Wavelength selection for portable noninvasive blood component measurement system based on spectral difference coefficient and dynamic spectrum. *Spectrochim Acta A Mol Biomol Spectrosc* 2018 Mar 15;193:40-46. [doi: [10.1016/j.saa.2017.10.063](https://doi.org/10.1016/j.saa.2017.10.063)] [Medline: [29223052](https://pubmed.ncbi.nlm.nih.gov/29223052/)]
133. Bobbia S, Macwan R, Benezeth Y, Mansouri A, Dubois J. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recogn Lett* 2019 Jun;124:82-90. [doi: [10.1016/j.patrec.2017.10.017](https://doi.org/10.1016/j.patrec.2017.10.017)]

134. Wijshoff RWCGR, Mischi M, Aarts RM. Reduction of Periodic Motion Artifacts in Photoplethysmography. *IEEE Trans Biomed Eng* 2017 Jan;64(1):196-207. [doi: [10.1109/tbme.2016.2553060](https://doi.org/10.1109/tbme.2016.2553060)]
135. Ram MR, Madhav KV, Krishna EH, Komalla NR, Reddy KA. A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Trans Instrum Meas* 2012 May;61(5):1445-1457. [doi: [10.1109/tim.2011.2175832](https://doi.org/10.1109/tim.2011.2175832)]
136. Chan K, Zhang Y. Adaptive reduction of motion artifact from photoplethysmographic recordings using a variable step-size LMS filter. 2002 Presented at: *Sensors 2002*; June 12-14, 2002; Orlando, Florida p. 1343-1346. [doi: [10.1109/icsens.2002.1037314](https://doi.org/10.1109/icsens.2002.1037314)]
137. Raghuram M, Madhav KV, Krishna EH, Komalla NR, Sivani K, Reddy KA. Dual-tree complex wavelet transform for motion artifact reduction of PPG signals. 2012 Presented at: *2012 IEEE international symposium on medical measurements and applications proceedings*; May 18-19, 2012; Budapest, Hungary p. 1-4. [doi: [10.1109/memea.2012.6226643](https://doi.org/10.1109/memea.2012.6226643)]
138. Jonathan E, Leahy M. Investigating a smartphone imaging unit for photoplethysmography. *Physiol Meas* 2010 Sep 24;31(11):N79-N83. [doi: [10.1088/0967-3334/31/11/n01](https://doi.org/10.1088/0967-3334/31/11/n01)]
139. Rendell M, Anderson E, Schlueter W, Mailliard J, Honigs D, Rosenthal R. Determination of hemoglobin levels in the finger using near infrared spectroscopy. *Clin Lab Haematol* 2003 Apr;25(2):93-97. [doi: [10.1046/j.1365-2257.2003.00493.x](https://doi.org/10.1046/j.1365-2257.2003.00493.x)] [Medline: [12641612](https://pubmed.ncbi.nlm.nih.gov/12641612/)]
140. Yim D, Baranoski GV, Kimmel B, Chen TF. A cell-based light interaction model for human blood. *Comput Graphic Forum* 2012;31:845-854. [doi: [10.1111/j.1467-8659.2012.03065.x](https://doi.org/10.1111/j.1467-8659.2012.03065.x)]
141. Jakovels D, Rubins U, Spigulis J. LASCA and PPG imaging for non-contact assessment of skin blood supply. In: *Medical Imaging 2013: Physics of Medical Imaging Proceedings volume 8668*.: International Society for Optics and Photonics; 2013 Presented at: *SPIE Medical Imaging 2013*; March 6, 2013; Orlando, Florida. [doi: [10.1117/12.2007375](https://doi.org/10.1117/12.2007375)]
142. Knorr-Chung BR, McGrath SP, Blike GT. Identifying airway obstructions using photoplethysmography (PPG). *J Clin Monit Comput* 2008 Apr 25;22(2):95-101. [doi: [10.1007/s10877-008-9110-7](https://doi.org/10.1007/s10877-008-9110-7)] [Medline: [18219579](https://pubmed.ncbi.nlm.nih.gov/18219579/)]
143. Ramasahayam S, Koppuravuri SH, Arora L, Chowdhury SR. Noninvasive blood glucose sensing using near infra-red spectroscopy and artificial neural networks based on inverse delayed function model of neuron. *J Med Syst* 2015 Jan 11;39(1):166. [doi: [10.1007/s10916-014-0166-2](https://doi.org/10.1007/s10916-014-0166-2)] [Medline: [25503416](https://pubmed.ncbi.nlm.nih.gov/25503416/)]
144. Smith AM, Mancini MC, Nie S. Bioimaging: second window for in vivo imaging. *Nat Nanotechnol* 2009 Nov;4(11):710-711 [FREE Full text] [doi: [10.1038/nnano.2009.326](https://doi.org/10.1038/nnano.2009.326)] [Medline: [19898521](https://pubmed.ncbi.nlm.nih.gov/19898521/)]
145. Yoon G, Lee JY, Jeon KJ, Park KK, Yeo HS, Hwang HT, et al. Identifying airway obstructions using photoplethysmography (PPG). In: *Proceedings Volume 4916, Optics in Health Care and Biomedical Optics: Diagnostics and Treatment*. 2002 Presented at: *Photonics Asia*; September 12, 2002; Shanghai, China. [doi: [10.1117/12.482947](https://doi.org/10.1117/12.482947)]
146. Yuan H, Memon SF, Newe T, Lewis E, Leen G. Motion artefact minimization from photoplethysmography based non-invasive hemoglobin sensor based on an envelope filtering algorithm. *Measurement* 2018 Feb;115:288-298. [doi: [10.1016/j.measurement.2017.10.060](https://doi.org/10.1016/j.measurement.2017.10.060)]
147. McDuff D, Gontarek S, Picard RW. Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera. *IEEE Trans Biomed Eng* 2014 Dec;61(12):2948-2954. [doi: [10.1109/tbme.2014.2340991](https://doi.org/10.1109/tbme.2014.2340991)]
148. Elgendi M. On the analysis of fingertip photoplethysmogram signals. *Curr Cardiol Rev* 2012 Feb 01;8(1):14-25 [FREE Full text] [doi: [10.2174/157340312801215782](https://doi.org/10.2174/157340312801215782)] [Medline: [22845812](https://pubmed.ncbi.nlm.nih.gov/22845812/)]
149. Nara S, Kaur M, Lal Verma K. Novel Notch Detection Algorithm for Detection of Dicrotic Notch in PPG Signals. *Int J Comput Appl* 2014 Jan 16;86(17):36-39. [doi: [10.5120/15081-3520](https://doi.org/10.5120/15081-3520)]
150. Shahrabaki SS, Ahmed B, Penzel T, Cvetkovic D. Photoplethysmography derivatives and pulse transit time in overnight blood pressure monitoring. 2016 Presented at: *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; Aug 16-20, 2016; Florida, USA p. 2855-2858. [doi: [10.1109/embc.2016.7591325](https://doi.org/10.1109/embc.2016.7591325)]
151. Chua C, Heneghan C. Continuous blood pressure monitoring using ECG and finger photoplethysmogram. 2006 Presented at: *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*; August 30 to September 3, 2006; New York, NY p. 5117-5120. [doi: [10.1109/iembs.2006.259612](https://doi.org/10.1109/iembs.2006.259612)]
152. Delle Chiaie L, Buck G, Grab D, Terinde R. Prediction of fetal anemia with Doppler measurement of the middle cerebral artery peak systolic velocity in pregnancies complicated by maternal blood group alloimmunization or parvovirus B19 infection. *Ultrasound Obstet Gynecol* 2001 Sep 12;18(3):232-236. [doi: [10.1046/j.0960-7692.2001.00540.x](https://doi.org/10.1046/j.0960-7692.2001.00540.x)] [Medline: [11555452](https://pubmed.ncbi.nlm.nih.gov/11555452/)]
153. Benaron DA, Benitz WE, Ariagno RL, Stevenson DK. Noninvasive methods for estimating in vivo oxygenation. *Clin Pediatr (Phila)* 1992 May 02;31(5):258-273. [doi: [10.1177/000992289203100501](https://doi.org/10.1177/000992289203100501)] [Medline: [1582091](https://pubmed.ncbi.nlm.nih.gov/1582091/)]
154. Flower RJ, Olsen RW, Van EMA. Apparatus and method for measuring blood constituents United States patent US 4,863,265. Google Patents. 1989. URL: <https://patents.google.com/patent/US9451920B2/en> [accessed 2021-03-18]
155. Poon C, Teng X, Wong Y, Zhang C, Zhang Y. Changes in the photoplethysmogram waveform after exercise. 2004 Presented at: *2nd IEEE/EMBS International Summer School on Medical Devices and Biosensors*; June 26 to July 2, 2004; Hong Kong, China p. 115-118. [doi: [10.1109/issmd.2004.1689576](https://doi.org/10.1109/issmd.2004.1689576)]
156. Menardi G, Torelli N. Training and assessing classification rules with imbalanced data. *Data Min Knowl Disc* 2012 Oct 30;28(1):92-122. [doi: [10.1007/s10618-012-0295-5](https://doi.org/10.1007/s10618-012-0295-5)]

157. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res* 2002 Jun 01;16:321-357. [doi: [10.1613/jair.953](https://doi.org/10.1613/jair.953)]
158. Talreja V, Ferrett T, Valenti MC, Ross A. Biometrics-as-a-service: A framework to promote innovative biometric recognition in the cloud. : January 12-15, 2018; 2018 Presented at: 2018 IEEE international conference on consumer electronics (ICCE); 2018; Las Vegas, USA p. 1-6. [doi: [10.1109/icce.2018.8326075](https://doi.org/10.1109/icce.2018.8326075)]
159. Hasan MK. BEst (Biomarker Estimation): health biomarker estimation non-invasively and ubiquitously, PhD Dissertation. Marquette University. 2019. URL: [https://epublications.marquette.edu/dissertations\\_mu/863/](https://epublications.marquette.edu/dissertations_mu/863/) [accessed 2021-03-17]

## Abbreviations

**AC:** alternating current  
**CBC:** complete blood count  
**CCD:** charge-coupled device  
**CMOS:** complementary metal oxide semiconductor  
**DC:** direct current  
**Hb:** hemoglobin  
**HbA<sub>1c</sub>:** glycated hemoglobin  
**ICA:** independent component analysis  
**InGaAs:** indium gallium arsenide  
**LED:** light-emitting diode  
**LOA:** limit of agreement  
**MAPE:** mean absolute percentage error  
**MLR:** multiple linear regression  
**MSE:** mean squared error  
**NIR:** near infrared  
**NIRS:** near infrared spectroscopy  
**PLSR:** partial least-squares regression  
**POC:** point of care  
**PPG:** photoplethysmography  
**RBC:** red blood cell  
**RGB:** red, blue, green  
**ROSE:** Random OverSampling Examples  
**SCD:** sickle cell diseases  
**SNR:** signal to noise ratio  
**SVR:** support vector machine regression

*Edited by G Eysenbach; submitted 26.10.19; peer-reviewed by G Dimauro, J Marcotte; comments to author 25.11.19; revised version received 20.01.20; accepted 10.02.20; published 08.04.21.*

*Please cite as:*

*Hasan MK, Aziz MH, Zarif MII, Hasan M, Hashem MMA, Guha S, Love RR, Ahamed S*

*Noninvasive Hemoglobin Level Prediction in a Mobile Phone Environment: State of the Art Review and Recommendations*

*JMIR Mhealth Uhealth* 2021;9(4):e16806

URL: <https://mhealth.jmir.org/2021/4/e16806>

doi: [10.2196/16806](https://doi.org/10.2196/16806)

PMID: [33830065](https://pubmed.ncbi.nlm.nih.gov/33830065/)

©Md Kamrul Hasan, Md Hasanul Aziz, Md Ishrak Islam Zarif, Mahmudul Hasan, MMA Hashem, Shion Guha, Richard R Love, Sheikh Ahamed. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 08.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Staff Perceptions of Preimplementation Barriers and Facilitators to a Mobile Health Antiretroviral Therapy Adherence Counseling Intervention in South Africa: Qualitative Study

Siobhan McCreesh-Toselli<sup>1\*</sup>, MSc; John Torline<sup>1</sup>, MD; Hetta Gouse<sup>1\*</sup>, PhD; Reuben N Robbins<sup>2\*</sup>, PhD; Claude A Mellins<sup>2\*</sup>, PhD; Robert H Remien<sup>2\*</sup>, PhD; Jessica Rowe<sup>3\*</sup>, MA; Neshaan Peton<sup>4\*</sup>, MPhil; Stephan Rabie<sup>1\*</sup>, PhD; John A Joska<sup>1\*</sup>, PhD

<sup>1</sup>HIV Mental Health Research Unit, Department of Psychiatry and Mental Health, University of Cape Town, Observatory, South Africa

<sup>2</sup>HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute and Columbia University, New York, NY, United States

<sup>3</sup>The Columbia Center for New Media Teaching and Learning, Columbia University, New York, NY, United States

<sup>4</sup>City of Cape Town Metropolitan Municipality, Cape Town, South Africa

\*these authors contributed equally

**Corresponding Author:**

John A Joska, PhD

HIV Mental Health Research Unit

Department of Psychiatry and Mental Health

University of Cape Town

J Block, Groote Schuur Hospital

Observatory

South Africa

Phone: 27 21 404 2174

Fax: 27 21 448 8158

Email: [john.joska@uct.ac.za](mailto:john.joska@uct.ac.za)

## Abstract

**Background:** South Africa adopted a universal test and treatment program for HIV infection in 2015. The standard of care that people living with HIV receive consists of 3 sessions of readiness counseling delivered by lay counselors (LCs). In the largest antiretroviral therapy (ART) program worldwide, effective and early HIV and ART education and support are key for ensuring ART adoption, adherence, and retention in care. Having LCs to deliver readiness counseling allows for the wide task-sharing of this critical activity but carries the risks of loss of standardization, incomplete content delivery, and inadequate monitoring and supervision. Systems for ensuring that a minimum standard of readiness counseling is delivered to the growing number of people living with HIV are essential in the care cascade. In resource-constrained, high-burden settings, mobile health (mHealth) apps may potentially offer solutions to these treatment gaps by providing content structure and delivery records.

**Objective:** This study aims to explore, at a large Cape Town-based nonprofit HIV care organization, the staff's perceived preimplementation barriers and facilitators of an mHealth intervention (Masivukeni) developed as a structured app for ART readiness counseling.

**Methods:** Masivukeni is a laptop-based app that incorporates written content, graphics, short video materials, and participant activities. In total, 20 participants were included in this study. To explore how an mHealth intervention might be adopted across different staff levels within the organization, we conducted 7 semistructured interviews (participants: 7/20, 35%) and 3 focus groups (participants in 2 focus groups: 4/20, 20%; participants in 1 focus group: 3/20, 15%) among LCs, supervisors, and their managers. In total, 20 participants were included in this study. Interviews lasted approximately 60 minutes, and focus groups ranged from 90 to 120 minutes. The Consolidated Framework for Implementation Research was used to explore the perceived implementation barriers and facilitators of the Masivukeni mHealth intervention.

**Results:** Several potential facilitators of Masivukeni were identified. Multimedia and visual elements were generally regarded as aids in content delivery. The interactive learning components were notably helpful, whereas facilitated updates to the adherence curriculum were important to facilitators and managers. The potential to capture administrative information regarding LC delivery and client logging was regarded as an attractive feature. Barriers to implementation included security risks and equipment costs,

the high volume of clients to be counseled, and variable computer literacy among LCs. There was uncertainty about the app's appeal to older clients.

**Conclusions:** mHealth apps, such as Masivukeni, were perceived as being well placed to address some of the needs of those who deliver ART adherence counseling in South Africa. However, the successful implementation of mHealth apps appeared to be dependent on overcoming certain barriers in this setting.

(*JMIR Mhealth Uhealth* 2021;9(4):e23280) doi:[10.2196/23280](https://doi.org/10.2196/23280)

## KEYWORDS

HIV/AIDS; antiretroviral treatment; low-resource settings; mHealth; Masivukeni; Consolidated Framework for Implementation Research; implementation research; lay antiretroviral therapy adherence counselors; mobile phone

## Introduction

### Background

In South Africa, the integration of lay counselors (LCs) in HIV health care has revealed some challenges in providing mental health services [1-3]. LCs face many structural challenges in their work, including little space and privacy to conduct counseling and limited support and supervision by health care staff [2,3]. To further compound these challenges, LCs are not part of the South African National Department of Health's formal employment structure but are rather supervised by their respective facility managers [3]. One of the challenges with integrating LCs in providing antiretroviral therapy (ART) adherence counseling includes the need to standardize content and delivery within a patient-centered framework [4,5]. Current research is focused on ways to optimize the skills of LCs who deliver these interventions [3]. A novel and growing avenue of research to address these challenges is using technology-based interventions [6-9]. Evidence suggests that the use of mobile health (mHealth) apps—mobile and wireless tools used in care settings—can positively impact postintervention health behaviors [10,11].

mHealth-based interventions offer several advantages, such as widespread access to professional care within resource-constrained environments, the incorporation of interactive components to enhance learning and medical education, and a variety of delivery modes (eg, via smartphones, tablets, and laptops) to suit various contexts [9,11]. Another benefit includes visual psychosocial education capabilities that accommodate different learning styles and literacy levels, which is important for lower literacy countries [7]. One such example is Masivukeni, a laptop-based mHealth intervention developed specifically for the South African context for use by LCs who deliver ART adherence counseling [7]. The intervention was informed by social action theory, which emphasizes social relationships and psychoeducation, crucial factors in positive adherence behavior [12]. The Masivukeni intervention provides a scaffold to deliver key information about HIV and ART and activities aimed at addressing behavioral factors that strengthen adherence [12]. Modules are presented in a linear fashion; they can be represented at any time and are delivered in a visual and interactive manner to improve patient interaction and reduce the effect of low literacy levels [7]. Interventions such as Masivukeni offer opportunities to improve access to standardized ART adherence counseling and enhance LCs' capacity to deliver interventions [13]. However, low-resource

health care settings present challenges to wider mHealth implementation because of issues such as limited funding, high costs, and poor infrastructure—factors known to impede the implementation process [10,14].

In South Africa, LCs provide some relief to the health care system but require substantial support, and mHealth solutions may be an effective tool to provide this support [1]. There are approximately 10,000 LCs in South Africa [15]. Although the selection criteria for LCs are highly variable, most LCs have completed secondary schooling [16], and in the Western Cape province, their training is provided by the People Development Center (PDC) [17]. In general, training includes (1) basic information on the HIV/AIDS disease process, (2) an overview of counseling and communication skills, (3) information on legal and ethical issues in HIV testing, (4) conducting pre- and postcounseling processes, and (5) record keeping and reporting practices [17]. Implementing interventions such as Masivukeni may provide standardized counseling, which is essential for ART program retention and patient well-being [2,3,6]. Research in this context is important for future implementation of mHealth apps to address the growing need to standardize content delivery and counseling in low-resource settings.

### Objectives

This study aims to explore staff perceptions of preimplementation barriers and facilitators of mHealth apps using Masivukeni.

## Methods

### Setting and Participants

The study was conducted across 4 sites in Cape Town, South Africa. All 4 sites were operated by TB/HIV Care, one of the largest nonprofit organizations involved in TB (tuberculosis) and HIV treatment and prevention in the Western Cape. TB/HIV Care largely supports local primary health care facilities by providing care to local communities.

Prospective participants were identified as those providing HIV testing services from TB/HIV Care. Participants and key stakeholders from this organization were therefore well placed to offer valuable insights into the perceived preimplementation barriers and facilitators of mHealth interventions, such as Masivukeni. Purposive sampling was used to recruit participants from the following groups: LCs, supervisors (area coordinators), and managers (district coordinators). At each respective site, 3-4 LCs, 1-2 supervisors, and 1-2 managers were recruited. LCs

are employed by TB/HIV Care but trained by the Western Cape Government, PDC, previously known as the AIDS Training Information Counseling Center. Participants were required to have at least one year of experience in their respective positions and have no previous exposure to Masivukeni.

A total of 3 focus groups (FGs) (participants in 2 FGs: 4/20, 20%; participants in 1 FG: 3/20, 15%) and 7 semistructured individual interviews (participants: 7/20, 35%) were included in the final analysis to assess perceptions of Masivukeni and mHealth interventions in general. A supervisor was excluded from the analysis because of familiarity with Masivukeni. In total, 16 women and 4 men participated in the study. At each site, we recruited 3-4 LCs, 1-2 supervisors, and 1-2 managers.

Data were collected by university-qualified senior research assistants and overseen by a master's level project manager. The senior research assistants were trained in qualitative data collection techniques and interviewing skills before data collection. The interviews and FGs were conducted in the participants' home language (ie, Afrikaans, English, or isiXhosa). The semistructured interviews' duration was approximately 60 minutes, whereas the duration of the FG discussions was 90 to 120 minutes. All interviews and FGs were audio recorded.

### Approach

The Consolidated Framework for Implementation Research (CFIR) was used to investigate the preliminary *barriers* and *facilitators* of Masivukeni. CFIR is a flexible framework that

incorporates constructs and strategies from various implementation theories and offers a collective approach toward implementation research [18-20]. It aims to address implementation at a multilevel system, which includes the intervention's characteristics, the staff who administer it, and the managers connected to the organization [18]. Determinant frameworks such as CFIR help to address factors hypothesized to influence implementation success [21]; the CFIR consists of 5 domains containing a number of constructs and subconstructs [18,21,22]. Domain 5, which is centered on the implementation process, was not included because this study focused on preimplementation. The interview guides were based on the overall themes of the CFIR domains. [Textbox 1](#) illustrates the 4 domains and an overview of the constructs. *Barriers* are defined as any factors perceived to potentially obstruct the implementation of Masivukeni. *Facilitators* are defined as any factors perceived to potentially enable the implementation of Masivukeni. *Neutral* is defined as a factor perceived as neither obstructing nor enabling the implementation of Masivukeni. To assess potential facilitators and barriers, participants were exposed to a 3-minute video demonstration of Masivukeni [23], followed by a discussion of the app and its potential use in their work environment.

Each domain consists of a set of constructs, whereas domain 3 also contains subconstructs. Domain 1 pertains to the characteristics of the intervention, whereas domains 2 and 3 refer to the inner and outer settings [18]. Domain 4 entails the characteristics of the stakeholders associated with and involved in the intervention [18].

**Textbox 1.** Domains 1-4 in the Consolidated Framework for Implementation Research.

## Domain 1: Intervention characteristics

- Innovation science
- Evidence strength and quality
- Relative advantage
- Adaptability
- Trialability
- Complexity
- Design quality and packaging
- Cost

## Domain 2: Outer setting

- Patient needs and resources
- Cosmopolitanism
- Peer pressure
- External policy and incentives

## Domain 3: Inner setting

- Structural characteristics
- Networks and communication
- Culture
- Implementation climate
- Organizational incentives and rewards
- Goals and feedback
- Learning climate
- Readiness for implementation

## Domain 4: Characteristics of the individuals

- Knowledge and beliefs about the innovation
- Self-efficacy
- Individual stage of change
- Individual identification with the organization
- Other personal attributes

**Data Management and Analysis**

Qualitative data were managed in Nvivo (QSR International, Version 11) and analyzed using the framework method [24] based on the CFIR domains. The framework method is a systematic qualitative analytical technique that produces highly structured outputs of summarized data [24]. For this study, a general deductive approach was applied, as the CFIR contains predefined codes listed as constructs. Constructs deemed significant to the study were determined by the a priori development of the CFIR model. As such, the data were applied to these constructs to determine the feasibility of the intervention. In addition, open-ended coding was included in the analysis to recognize commonalities and peculiarities within the data for the potential development of constructs not listed in the framework. A codebook template from the CFIR database

with construct definitions was used [25]. Minor modifications were made to the constructs to fit the context of the study. For example, in domain 4, there is a construct called *other personal attributes*. This construct was broadly defined to include other personal features such as competency, motivation, and values, which can relate to the individuals themselves and to the intervention [25]. For our study's purpose, this construct was refined to include personally identified challenges, which specified challenges that LCs, supervisors, and managers experienced in their day-to-day work routine.

Two analysts used the same codebook to analyze the data independently. A case report was generated for each FG and individual interviews organized into CFIR domains and constructs. Case reports contained summary statements to illustrate how constructs manifested and were supported by



excerpts from raw data sets. Upon completion of analysis, all case reports were compared to identify patterns and differences among the stakeholders. To identify barriers and facilitators, the constructs were subjected to a rating procedure [26]. The rating was based on the valence and strength of the statements under each construct [27]. Valence related to constructs is thought to have a positive or negative influence on implementation. These statements can be positive, negative, neutral, or mixed [27]. A mixed statement contains an equal number of positive and negative statements. In terms of strength, statements may be either strongly (2) or weakly (1) positive or negative. To determine the strength of statements, analysts considered several factors, including the level of agreement between respondents, their use of specific and concrete examples, and choice of language [27]. Any discrepancies in coding were overviewed by a third analyst and discussed with the independent analysts until consensus on the reports was reached.

### Ethical Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the University of Cape Town Human Research Ethics Committee (reference number: 246/2011). All participants in the study participated freely and voluntarily and provided written informed consent.

## Results

### Overview

The following section includes CFIR constructs that were significantly positive or negative among participants during the analysis. Constructs that were not significant to the study or contained no data were excluded. Although open-ended coding for potential inductively developed constructs was included, the analyzed data were easily placed into the predefined constructs and did not necessitate further development of new constructs.

### Domain 1: Intervention Characteristics

#### Facilitator: Perceived Evidence Strength and Quality

Facilitator is a supporting belief that the intervention will achieve desired outcomes. The LCs, supervisors, and managers all felt that Masivukeni would be useful to counselors and their clients. This was especially true when transferring knowledge and facilitating psychosocial education in a standardized manner:

*...it's visual and it's something that will stick and you will have it right in front of you so you won't forget information or leave things out that are important because every step is right there...* [Manager]

#### Facilitator: Relative Advantage

Compared with the current model used in adherence counseling, the LCs, supervisors, and facility managers believed that the mHealth app had several advantages over their existing model, in which flipcharts are used as a supportive visual aid. First, the interactive nature and the visually stimulating components of the activities were thought to be more engaging. Second, all 3 groups of participants felt that this mHealth intervention would

be particularly helpful when working with adolescents, especially because they may be attracted to the use of technology. The same advantage for older clientele was not clear, and some LCs mentioned that they would like to keep the books and flipcharts for the grannies:

*...we thought this is something that sounds like it is going to be more helpful than what we are doing, and we are doing as much as we can with talks, flipcharts and all. But we do notice that the clients, they are different...some they understand well, some they understand with pictures...we thought this is a very good idea because...it is like both of you are involved in this learning activity.* [Supervisor]

The third major advantage was specific to supervisors and managers. They felt that having an mHealth app may potentially assist with their heavy administrative workload. Currently, Masivukeni is able to record some features of counseling sessions, such as session length. The participants were excited about the sharing and monitoring capabilities of the app, especially the ability to check the number of patients seen per day by each logged-in LC:

*...I can easily pull up the stats for the day and see, okay, this particular counsellor has seen these particular clients for this day. ... I don't have to wait for the month end whereby I'm going to get the facility diary and check everything.* [Manager]

#### Facilitator: Recommendations for Adapting the Intervention to Meet Local Needs

Many recommendations were made to potentially adapt Masivukeni. The LCs felt that the app should have alternate language options available, specifically isiXhosa and Afrikaans—common regional languages in the context of the study. Although counseling sessions are conducted in English, LCs revealed that educational explanations often occurred in the client's first language, especially if the client was older and less proficient in English. In addition, the participants from each group felt that using an app such as Masivukeni for children, adolescents, adults, and older people was possible but that the content and visual components might need changes for different levels of understanding and appropriateness depending on the age group. The supervisors requested that more information on nutrition and activities aimed at identifying goals and barriers be added to the app. Information on nutrition and identification of barriers were not presented in the demonstration:

*...you need to understand also not everybody is perfect in English. You need to understand also there are youngsters, there are "Gogos" [Grandmothers in isiXhosa] of 50 years old. They don't understand all these things. So, we need to present it in a manner that will suit everyone's understanding.* [Counselor]

LCs are required to discuss sensitive topics with patients, such as male circumcision, which can be considered taboo for women to speak about in some cultures. Although LCs are trained in approaching these topics, the app could include activities on sensitive discussions to ease difficulties; this was thought to be potentially helpful.

Other suggested modifications to the intervention included questions regarding its potential electronic capabilities. Specifically, queries centered on whether this type of app could include automated software updates to access the latest adherence curricula. This concern was present across all 3 groups. The participants expressed a need to gain access to changing adherence information and treatment of HIV. Other smaller electronic modifications, such as space for electronic signatures on consent forms, were appealing features. Finally, the groups of participants felt that Masivukeni would be useful as a training tool at the PDC or newly appointed ART LCs in the field, and not just as an aid during ART adherence counseling:

*We have to wait for ATICC [PDC] to give us dates so that we can send them [the LCs] for updates...But now, if it can also automatically update the information when there's something that has been approved so that it can also show up, then it will be also helpful. [Manager]*

### **Barrier: Complexity to Use the Intervention and Implement It in the Current Setting**

There was slight disagreement among the participants in terms of whether the LCs would be able to use the app. Computer literacy among LCs is a concern. The participants revealed that some of this concern was specific to older LCs, who were not familiar with or open to using computers and technology. However, the facility managers felt that using technology would not be a problem, especially if smartphones or tablets were the modes of delivery instead of laptops:

*...most of them [are] already using smart phones. It's only two or three that don't have smart phones and they already know because a tab[let] is actually similar to a smart phone...They wouldn't struggle with that at all from my point of view. [Manager]*

*It's not easy to use computers...even to use my tablet. I do have the tablet but sometimes I'm stuck. [Counselor]*

### **Neutral: Perceived and Associated Costs**

The supervisors and facility managers presented some concern about which entity would, on the one hand, purchase the equipment, and on the other hand, provide its maintenance:

*... who's going to be responsible for the equipment and who is going to buy it? [Supervisor]*

However, there was some recognition that the maintenance of this app, compared with the upkeep of physical materials such as flipcharts and books, may prove more cost-effective:

*I feel it's easier to change this application than to have all those new posters and things being printed and it tears and it breaks... [Manager]*

## **Domain 2: Outer Setting**

### **Facilitator: Perceived Needs and Resources of the Clients**

A clear focus on the needs of clients and patients from relevant stakeholders plays a positive role in the effective implementation

of an intervention [18,19]. Furthermore, the ability of an intervention to meet these needs is paramount. The clients' perceived needs centered on access to important information and education on HIV/TB and access to ART adherence education and counseling, which Masivukeni is designed to achieve. The participants felt that addressing barriers to adherence and, in particular, the issue of disclosure was a recurring difficulty among clients. Another important need focused on the issue of accessibility to clinics and facilities. Specifically, clients sometimes struggle to find transport to and from clinics because of financial restraints. When counselors are aware that a client will not be able to return, LCs sometimes combine 2 counseling sessions into 1 session. Thus, the degree of flexibility for each client is exercised by LCs. How Masivukeni addresses issues of accessibility requires further adaptation. Finally, LCs mentioned that clients sometimes migrate to different clinics and facilities around the Western Cape, repeating HIV testing and counseling with various LCs. Masivukeni, as an mHealth app, is currently able to record basic session parameters and may provide a digital space to log information:

*Some of the clients, they default because they didn't disclose at home, and then they give us the wrong addresses...Others, they came, but they work too much, so they don't have a chance to come to the clinics. [Counselor]*

*When we go to the community, we will find out that [a client]...has got a folder at each and every clinic. [Counselor]*

### **Neutral: Peer Pressure**

Although not a significantly positive or negative construct, there is an awareness of the use of technological interventions in health care in South Africa. Some LCs mentioned MomConnect and asked whether the Masivukeni app was similar to it. The LCs felt that the older LCs and clientele may not respond to technological interventions:

*I take it from the MomConnect. There is this web that Government also has, where you connect pregnant mummies onto the site. The only ones that connect to the site is the young ones. You get the 39-year-olds, the 40-year-olds...we want them to connect to the site so that they can see this is danger signs in pregnancy, this is what can go wrong and all of that. But they are not technology orientated. [Counselor]*

### **Barrier: External Policy and Incentives—Reaching Clients' Daily Targets**

The LCs, supervisors, and facility managers expressed the concern that reaching the required daily targets of clients per day is overwhelming and a significant stressor. Masivukeni contains a structured platform and takes a certain amount of time to complete each activity. There was significant apprehension as to whether the intervention would take more time than currently exercised per client, consequently hindering the daily target attainment:

*You have to reach the target; that is why you have to rush the time. [Counselor]*

### Domain 3: Inner Setting

#### **Neutral: Networks and Communications—The Quality and Nature of Relationships**

LCs sometimes feel undermined by doctors and nurses. The LCs mentioned that they are often questioned about clients' whereabouts when the facility is struggling to make contact with clients. A number of factors contribute to difficulties tracking down clients, such as frequent changes in cell phone numbers and client migration. Some LCs reported that this contributed to feelings of stress as they felt responsible for their clients' whereabouts. However, the relationship between LCs and supervisors was positive. For example, LCs reported that they supported each other during difficult cases.

#### **Facilitator: Cultural Values—The Perceived Role of LCs**

The role of LCs centered on providing a supportive relationship to their clients, with the aim to remain adherent to ART treatment. Masivukeni, described as an mHealth app used by LCs to facilitate ART adherence [7], may help to foster these relationships and support adherence counseling. Supervisors and managers added that LCs endeavor to empower their clients and help them address barriers to adherence. In this regard, Masivukeni contains specific content to address these barriers [7]. The facility managers agreed that a major aim of the organization is to reach the daily target of clients seen per day, in addition to providing standardized, quality counseling. Although Masivukeni cannot facilitate reaching the daily target, it can facilitate standardized adherence counseling. These organizational values and norms can be broadly aligned with the aims and content of Masivukeni.

#### **Facilitator: Implementation Climate as a Shared Receptivity for Change**

The LCs, supervisors, and facility managers expressed their openness to the implementation of Masivukeni, although there was some reservation about whether this would interfere with time pressures in their daily duties. This rating reflects the summation of the following subconstructs:

##### **Tension for Change**

An mHealth solution was needed in the following 3 specific areas: an electronic database to (1) assist with administration and the daily workload; (2) streamline counseling and testing processes; and (3) provide access to evolving medical information and treatments for HIV and ART. A manager stated:

*Some facilities have been asking for flip charts now for some time and with the flip charts, things change, then we ask when are we going to update it because the information changes...*

##### **Compatibility of the Intervention With Existing Systems and Workflow**

Differences between LCs and their supervisors and managers were evident in this construct. The LCs and two supervisors seemed concerned that the app would interrupt the workflow

of LCs and stop them from completing the daily targets. This is because Masivukeni is highly structured and takes a set amount of time to complete each session. LCs expressed the need to remain flexible for each client's needs, which can result in different amounts of time spent with their clients. The LCs and a supervisor specifically suggested that this app would be useful at their clubs, where time is not as much of a constraint. However, the facility managers felt that having an app like this for the clubs and individual sessions was useful:

*I am thinking this can be used in their groups, clubs. It will be much better to use it in the group. It is –One-on-one might take very long...we can identify the clients that are already having a problem of defaulting or having challenges with their adherence. [Supervisor]*

LCs also indicated that laptops pose a security risk. Laptops were seen as valuables, and the facilities had been previously burgled. Security, although mentioned, was less of a concern if the app was delivered on a smaller device (such as a smartphone or tablet computer equivalent), which they thought may easily be locked away and kept safe:

*...at our clinic, we have got a problem of burglary. They have burgled that container three times. Our chairs like this were gone. The doctor's computer was gone. [Counselor]*

##### **Relative Priority to Implement**

There were some differences between the groups. LCs and many supervisors did not prioritize the implementation of Masivukeni. However, the facility managers expressed that implementing an mHealth app such as Masivukeni, with features that allow for an electronic database, is necessary. A facility manager mentioned that standardization of counseling sessions is paramount to ensure quality service and that if Masivukeni could provide this, it would be important to begin implementation.

##### **Learning Climate to Learn a New Method**

Despite concerns about time constraints and the apparent rigidity of the app, LCs, supervisors, and facility managers showed a desire to learn to use technology in their field. The participants strongly felt that proper training programs and an ongoing technology support team were necessary for effective implementation:

*I think innovation is good. We use technology here now these days. To do something we have got apps, different apps...It will also help the health side if things like this can be developed so that it could be easier...Not only for the user but also for the client. [Counselor]*

##### **Neutral: Readiness for Implementation**

There was not enough data for these constructs to be significantly positive or negative. However, 2 subconstructs—leadership engagement and available resources—were manifested among the participants.

### **Leadership Engagement for Future Implementation Endeavors**

The supervisors and facility managers expressed that for effective implementation to occur, staff engagement would be necessary, particularly related to training and motivation. Supervisors and facility managers expressed a desire to involve the LCs during the implementation process, collaborate with researchers on further adaptations to the app, and develop implementation strategies that will best fit the organization:

*I think if myself, as being the supervisor, is positive about the thing that we try out, then it will be a step in the right direction. [Supervisor]*

### **Available Resources That Can Include Physical Space and Allotted Time**

LCs and supervisors mentioned that there were issues with adequate space management. Specific issues are that LCs share consultation rooms, which they feel compromises confidentiality and privacy during their sessions. Use, safekeeping, and maintenance of the equipment (ie, the Masivukeni laptop) were of some concern for future implementation.

### **Domain 4: Individuals Characteristics**

#### ***Facilitator: Knowledge and Beliefs About the Intervention***

General sentiments regarding Masivukeni were positive. Participants felt that it may assist LCs by providing counseling and facilitating psychosocial education for clients. However, again, the participants were unsure and concerned about how the structured and step-by-step nature of the app would fit their daily duties. Although not indicated by the researchers or demonstrated in the video, participants projected 2 capabilities onto Masivukeni. The first was electronic database recording for clients that included a sharing feature and internet access to a web-based system.

#### ***Facilitator: Personally Identified Challenges***

LCs mentioned that some difficulties include dealing with an array of complex psychosocial issues, such as HIV-positive pregnant teenagers. The LCs, supervisors, and facility managers indicated that disclosure and clients who default are also significant stressors to LC work. In turn, supervisors and facility managers expressed concerns that LCs are not always prepared with sufficient skills to deal with the variety of difficulties each client presents. Masivukeni may be able to support LCs when dealing with complex psychosocial issues.

## ***Discussion***

### **Principal Findings**

This study explored the perceived preimplementation barriers and facilitators of an mHealth HIV/ART counseling intervention, Masivukeni, in a South African setting. As technology is used more widely in the public health sphere, particularly in low-resource settings, there is a need to assess effective ways to ensure reliable and sustainable scale-up strategies for mHealth interventions [28-30]. The study revealed general, shared receptivity for the inclusion of technological

apps such as Masivukeni in clinics and health care facilities. This shared receptivity is critical for effective implementation [31]. Potential end users of Masivukeni felt that there were several advantages; these entailed (1) the ability to include key counseling techniques and materials to overcome health challenges experienced by patients (such as problem-solving activities), (2) accessibility to subpopulations (such as adolescents) because of the interactive e-environment, (3) the potential for a rapid, real-time update of materials, and (4) the possibility of enhancing the record keeping and monitoring functions required by providers.

LCs, supervisors, and facility managers were also excited about the prospect of using Masivukeni to enhance staff skills. Previous research has shown that LCs trained to use Masivukeni expressed feelings of empowerment [32], which is useful as LCs deal with a variety of complex cases that can cause significant emotional stress [33]. An app that facilitates basic counseling skills and knowledge through interactive and engaging learning can be impactful [11]. Participants felt that Masivukeni would be able to accomplish knowledge transfer and may also help to engage with adolescent subpopulations, who they thought required more attention-grabbing material. There is some evidence suggesting that adolescents and young adults may benefit from tailored, interactive interventions [34]; thus, using mHealth interventions tailored for adolescents may be particularly useful.

Masivukeni can record certain counseling session parameters, such as session length and which psycho-educational topics were covered. It also contains a simple referral system. Adaptations to include a basic electronic record (eRecord) database are possible and should be considered. Initial research on eRecord use in low-resource settings has shown that it may help relieve vast amounts of paperwork [35,36]. eRecord was appealing because it would ease (1) logging client information to help with migration issues and (2) monitoring counseling sessions.

Three key suggestions were noted. First, participants suggested smartphones and tablets would be preferred over computers or laptops as the mode of intervention delivery because LCs are more familiar with that technology, requiring less training. Recent evidence suggests that mobile technology access is on the rise in low-resource settings [28] and may be a better option. In addition, smartphones and tablets, which were perceived as smaller and less costly devices, were also thought to be less of a security threat. The second suggestion entailed the inclusion of several language options, as different languages and various computer literacy levels exist in South Africa [37]. Different literacy levels have been found in other resource-limited settings [37,38]. Training workshops to address computer literacy are needed for successful implementation. Finally, many LCs and supervisors were concerned that Masivukeni's structured platform might become a significant barrier because of the limited time each LC has with their clients. LCs and supervisors suggested that this app could be used in group therapy wherein time is a significant constraint factor, as opposed to one-on-one sessions. However, if the mHealth app was adapted to fit within each session's time constraints, this barrier may be less significant. Meeting these adaptations will be important because

interventions that pose a perceived threat may evoke resistance to intervention adoption [30].

There are 3 major barriers. The first barrier concerns which entity would be responsible for purchasing costs and maintaining mHealth interventions such as Masivukeni. This is important because a key advantage of using mHealth apps in low-resource settings is their potential to be cost-effective [1,7]. Potential technical support may also become a necessity if staff are not technology-oriented. The second barrier deals with the sentiment that LCs may, at times, feel undermined by other staff members [39]. This study pertained to communication difficulties between doctors and LCs concerning LC roles and responsibilities; this has been reported in previous studies [3,9]. Successful intervention implementation depends on effective communication between staff members [18] and can be negatively affected in low-resource environments because of unclear health system responsibilities [14]. In this case, unclear career paths for LCs and poor communication may play a role in the future success of any intervention implementation, not just mHealth apps. Finally, the LCs revealed that their counseling space was sometimes inadequate and not always private, compromising on confidentiality. Although this barrier may not be specific to Masivukeni, lack of available resources, such as physical space, may hinder the potential implementation of interventions in general [6,7,13]. However, as an example of a tailored mHealth intervention, Masivukeni may be able to fill the need for standardized ART adherence practices [2,3]. Despite these barriers, the implementation climate was open and optimistic.

### Limitations

The demonstration of Masivukeni did not show its latest features and modifications. For example, alternative language options were a concern among the participants. However, later versions of Masivukeni include activities in isiXhosa (the predominant local Bantu language in the Western Cape region). In addition, because the participants were not familiar with the entirety of the intervention and were only exposed to a demonstrational video, the perceptions represented in the study were only an initial indication of the implementation context. However, an investigation into the perceptions of mHealth solutions and existing knowledge around such interventions are informative steps in the direction of intervention adoption and dissemination

[21]. Another limitation of this study is the lack of external validity because of the limited number of participants in the sample. However, the in-depth nature of the information collected in this study aimed to address this limitation.

### Conclusions and Future Implications

This study qualitatively explored staff perceptions of preimplementation barriers and facilitators of Masivukeni, an mHealth app. Participants were excited about the potential use of Masivukeni in their workplace and felt that it had several advantages. Opportunities to implement the tool include first including it in the adherence counseling curriculum offered at the local training center. This affords the future counselor the benefit of building the theory and practice of counseling into the platform provided, familiarity with the tool, and mastery over its use, which may lead to less time taken to administer the modules and also ease concerns about lack of skills to use the app. A second implementation recommendation would be to allow the use of the tool in selected clinics in a piecemeal fashion. Hopefully, this would reduce anxieties around the time taken to administer the full intervention and seed out its use without insisting on it in its entirety. These strategies may not only facilitate implementation but also reduce perceived barriers. Participants made various suggestions about the ways in which the app could be adapted, which may show some degree of staff motivation and engagement. However, there were concerns about the compatibility with the LC workflow and load, which may contribute to the app's poor longevity [40].

Questions regarding the full-scale implementation of apps such as Masivukeni remain. First, where would the app be best suited? LCs and some supervisors felt it would be best suited for group therapy, whereas other supervisors and facility managers felt it would be of best use in individual sessions. Second, how do older LCs and clients, who are not technology-oriented, react to technology use? Can interventions such as Masivukeni be adapted to reach those clients who have transportation challenges? To this end, counseling sessions may be delivered via smartphones or tablets as a possible future adaptation. Future studies should also look at the suitability of mHealth apps such as Masivukeni for adolescents; our study participants felt this app would be particularly useful and engaging for this subpopulation.

---

### Acknowledgments

The authors would like to extend thanks to Molemoeng Shebi-Magadla and all TB/HIV Care staff members for their assistance. The authors would also like to thank the participants for their contributions and involvement in this study. This study was supported by the following 2 National Institute of Mental Health grants: R01-MH95576-05 (principal investigator: RHR) and P30-MH43520 (principal investigator: RHR).

---

### Conflicts of Interest

None declared.

---

### References

1. Callaghan M, Ford N, Schneider H. A systematic review of task- shifting for HIV treatment and care in Africa. *Hum Resour Health* 2010 Mar 31;8:8. [doi: [10.1186/1478-4491-8-8](https://doi.org/10.1186/1478-4491-8-8)] [Medline: [20356363](https://pubmed.ncbi.nlm.nih.gov/20356363/)]

2. Dewing S, Mathews C, Cloete A, Schaay N, Shah M, Simbayi L, et al. From research to practice: lay adherence counsellors' fidelity to an evidence-based intervention for promoting adherence to antiretroviral treatment in the Western cape, South Africa. *AIDS Behav* 2013 Nov;17(9):2935-2945. [doi: [10.1007/s10461-013-0509-x](https://doi.org/10.1007/s10461-013-0509-x)] [Medline: [23666183](https://pubmed.ncbi.nlm.nih.gov/23666183/)]
3. Petersen I, Fairall L, Egbe CO, Bhana A. Optimizing lay counsellor services for chronic care in South Africa: a qualitative systematic review. *Patient Educ Couns* 2014 May;95(2):201-210 [FREE Full text] [doi: [10.1016/j.pec.2014.02.001](https://doi.org/10.1016/j.pec.2014.02.001)] [Medline: [24629835](https://pubmed.ncbi.nlm.nih.gov/24629835/)]
4. Dewing S, Mathews C, Schaay N, Cloete A, Louw J, Simbayi L. Behaviour change counselling for ARV adherence support within primary health care facilities in the Western Cape, South Africa. *AIDS Behav* 2012 Jul;16(5):1286-1294. [doi: [10.1007/s10461-011-0059-z](https://doi.org/10.1007/s10461-011-0059-z)] [Medline: [21983698](https://pubmed.ncbi.nlm.nih.gov/21983698/)]
5. Dewing S, Mathews C, Schaay N, Cloete A, Louw J, Simbayi L. "It's Important to Take Your Medication Everyday Okay?" An Evaluation of Counselling by Lay Counsellors for ARV Adherence Support in the Western Cape, South Africa. *AIDS Behav* 2012 May 19;17(1):203-212. [doi: [10.1007/s10461-012-0211-4](https://doi.org/10.1007/s10461-012-0211-4)]
6. Robbins RN, Gouse H, Brown HG, Ehlers A, Scott TM, Leu C, et al. A Mobile App to Screen for Neurocognitive Impairment: Preliminary Validation of NeuroScreen Among HIV-Infected South African Adults. *JMIR Mhealth Uhealth* 2018 Jan 05;6(1):e5. [doi: [10.2196/mhealth.9148](https://doi.org/10.2196/mhealth.9148)]
7. Remien RH, Mellins CA, Robbins RN, Kelsey R, Rowe J, Warne P, et al. Masivukeni: development of a multimedia based antiretroviral therapy adherence intervention for counselors and patients in South Africa. *AIDS Behav* 2013 Jul;17(6):1979-1991 [FREE Full text] [doi: [10.1007/s10461-013-0438-8](https://doi.org/10.1007/s10461-013-0438-8)] [Medline: [23468079](https://pubmed.ncbi.nlm.nih.gov/23468079/)]
8. Sinclair PM, Kable A, Levett-Jones T, Booth D. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: A systematic review. *Int J Nurs Stud* 2016 May;57:70-81. [doi: [10.1016/j.ijnurstu.2016.01.011](https://doi.org/10.1016/j.ijnurstu.2016.01.011)] [Medline: [27045566](https://pubmed.ncbi.nlm.nih.gov/27045566/)]
9. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P T* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](https://pubmed.ncbi.nlm.nih.gov/24883008/)]
10. Claborn KR, Fernandez A, Wray T, Ramsey S. Computer-based HIV adherence promotion interventions: a systematic review: Translation Behavioral Medicine. *Transl Behav Med* 2015 Sep;5(3):294-306 [FREE Full text] [doi: [10.1007/s13142-015-0317-0](https://doi.org/10.1007/s13142-015-0317-0)] [Medline: [26327935](https://pubmed.ncbi.nlm.nih.gov/26327935/)]
11. Lustria MLA, Noar SM, Cortese J, Van SSK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013;18(9):1039-1069. [doi: [10.1080/10810730.2013.768727](https://doi.org/10.1080/10810730.2013.768727)] [Medline: [23750972](https://pubmed.ncbi.nlm.nih.gov/23750972/)]
12. Ewart CK. Social action theory for a public health psychology. *Am Psychol* 1991 Sep;46(9):931-946. [doi: [10.1037//0003-066x.46.9.931](https://doi.org/10.1037//0003-066x.46.9.931)] [Medline: [1958012](https://pubmed.ncbi.nlm.nih.gov/1958012/)]
13. Robbins RN, Mellins CA, Leu C, Rowe J, Warne P, Abrams EJ, et al. Enhancing Lay Counselor Capacity to Improve Patient Outcomes with Multimedia Technology. *AIDS Behav* 2015 Jun;19 Suppl 2:163-176. [doi: [10.1007/s10461-014-0988-4](https://doi.org/10.1007/s10461-014-0988-4)] [Medline: [25566763](https://pubmed.ncbi.nlm.nih.gov/25566763/)]
14. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC Public Health* 2014 Feb 21;14(1). [doi: [10.1186/1471-2458-14-188](https://doi.org/10.1186/1471-2458-14-188)]
15. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. *Disabil Rehabil* 2012;34(19):1608-1616. [doi: [10.3109/09638288.2012.656792](https://doi.org/10.3109/09638288.2012.656792)] [Medline: [22489612](https://pubmed.ncbi.nlm.nih.gov/22489612/)]
16. Kagee A. Training lay counsellors in public health: Considerations for social workers, professional counsellors and psychologists. *Global Public Health* 2020 Feb 21;15(6):918-924. [doi: [10.1080/17441692.2020.1730931](https://doi.org/10.1080/17441692.2020.1730931)]
17. Western CG. People Development Centre (PDC) Online website. 2020. URL: [https://www.westerncape.gov.za/your\\_gov/185](https://www.westerncape.gov.za/your_gov/185)
18. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009;4:50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
19. Damschroder LJ, Lowery JC. Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR). *Implement Sci* 2013;8:51 [FREE Full text] [doi: [10.1186/1748-5908-8-51](https://doi.org/10.1186/1748-5908-8-51)] [Medline: [23663819](https://pubmed.ncbi.nlm.nih.gov/23663819/)]
20. Keith RE, Crosson JC, O'Malley AS, Crompton D, Taylor EF. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. *Implement Sci* 2017 Dec 10;12(1):15 [FREE Full text] [doi: [10.1186/s13012-017-0550-7](https://doi.org/10.1186/s13012-017-0550-7)] [Medline: [28187747](https://pubmed.ncbi.nlm.nih.gov/28187747/)]
21. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci* 2015;10:53 [FREE Full text] [doi: [10.1186/s13012-015-0242-0](https://doi.org/10.1186/s13012-015-0242-0)] [Medline: [25895742](https://pubmed.ncbi.nlm.nih.gov/25895742/)]
22. Smith LR, Damschroder L, Lewis CC, Weiner B. The Consolidated Framework for Implementation Research: advancing implementation science through real-world applications, adaptations, and measurement. *Implementation Sci* 2015 Aug 14;10(S1). [doi: [10.1186/1748-5908-10-s1-a11](https://doi.org/10.1186/1748-5908-10-s1-a11)]

23. ColumbiaLearn. A demonstration of Masivukeni. -08-29. 2012. URL: <https://www.youtube.com/watch?v=MOGlicmko0> [accessed 2018-01-19] [WebCite Cache ID 6waQLePWX]
24. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [FREE Full text] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
25. Codebook Template. CFIR Technical Assistance Website. URL: <http://www.cfirguide.org/tools.html> [accessed 2018-01-19] [WebCite Cache ID 6waVeRgUX]
26. Qualitative Data: Data Analysis. CFIR Technical Assistance Website-29-10. 2014. URL: <http://www.cfirguide.org/qual.html> [accessed 2018-01-24] [WebCite Cache ID 6whsjlgz5]
27. Qualitative Data: Data Analysis, Rating Rules. CFIR Technical Assistance Website-29-10. 2014. URL: <http://www.cfirguide.org/RatingRules10.29.14.pdf> [accessed 2018-02-19] [WebCite Cache ID 6xLPI6379]
28. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Med* 2013;10(2):e1001382 [FREE Full text] [doi: [10.1371/journal.pmed.1001382](https://doi.org/10.1371/journal.pmed.1001382)] [Medline: [23424286](https://pubmed.ncbi.nlm.nih.gov/23424286/)]
29. English M. Designing a theory-informed, contextually appropriate intervention strategy to improve delivery of paediatric services in Kenyan hospitals. *Implementation Sci* 2013 Mar 28;8(1). [doi: [10.1186/1748-5908-8-39](https://doi.org/10.1186/1748-5908-8-39)]
30. Barker PM, Reid A, Schall MW. A framework for scaling up health interventions: lessons from large-scale improvement initiatives in Africa. *Implementation Sci* 2016 Jan 29;11(1). [doi: [10.1186/s13012-016-0374-x](https://doi.org/10.1186/s13012-016-0374-x)] [Medline: [26821910](https://pubmed.ncbi.nlm.nih.gov/26821910/)]
31. Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009 Oct 19;4:67 [FREE Full text] [doi: [10.1186/1748-5908-4-67](https://doi.org/10.1186/1748-5908-4-67)] [Medline: [19840381](https://pubmed.ncbi.nlm.nih.gov/19840381/)]
32. Gouse H, Robbins R, Mellins, Kingon A, Rowe J, Henry M, et al. Empowering Lay-Counsellors with Technology: Masivukeni, a Standardized Multimedia Counselling Support Tool to Deliver ART Counselling. *AIDS Behav* 2018 Oct;22(10):3345-3356 [FREE Full text] [doi: [10.1007/s10461-018-2145-y](https://doi.org/10.1007/s10461-018-2145-y)] [Medline: [29779161](https://pubmed.ncbi.nlm.nih.gov/29779161/)]
33. Tucker JD, Tso LS, Hall B, Ma Q, Beanland R, Best J, et al. Enhancing Public Health HIV Interventions: A Qualitative Meta-Synthesis and Systematic Review of Studies to Improve Linkage to Care, Adherence, and Retention. *EBioMedicine* 2017 Mar;17:163-171 [FREE Full text] [doi: [10.1016/j.ebiom.2017.01.036](https://doi.org/10.1016/j.ebiom.2017.01.036)] [Medline: [28161401](https://pubmed.ncbi.nlm.nih.gov/28161401/)]
34. Naar-King S, Outlaw AY, Sarr M, Parsons JT, Belzer M, Macdonell K, Adolescent Medicine Network for HIV/AIDS Interventions. Motivational Enhancement System for Adherence (MESA): pilot randomized trial of a brief computer-delivered prevention intervention for youth initiating antiretroviral treatment. *J Pediatr Psychol* 2013 Jul;38(6):638-648 [FREE Full text] [doi: [10.1093/jpepsy/jss132](https://doi.org/10.1093/jpepsy/jss132)] [Medline: [23359664](https://pubmed.ncbi.nlm.nih.gov/23359664/)]
35. Zargarani E, Schuurman N, Nicol AJ, Matzopoulos R, Cinnamon J, Taulu T, et al. The electronic Trauma Health Record: design and usability of a novel tablet-based tool for trauma care and injury surveillance in low resource settings. *J Am Coll Surg* 2014 Jan;218(1):41-50. [doi: [10.1016/j.jamcollsurg.2013.10.001](https://doi.org/10.1016/j.jamcollsurg.2013.10.001)] [Medline: [24355875](https://pubmed.ncbi.nlm.nih.gov/24355875/)]
36. Bull S, Thomas DS, Nyanza EC, Ngallaba SE. Tanzania Health Information Technology (T-HIT) System: Pilot Test of a Tablet-Based System to Improve Prevention of Mother-to-Child Transmission of HIV. *JMIR Mhealth Uhealth* 2018 Jan 15;6(1):e16 [FREE Full text] [doi: [10.2196/mhealth.8513](https://doi.org/10.2196/mhealth.8513)] [Medline: [29335236](https://pubmed.ncbi.nlm.nih.gov/29335236/)]
37. Sukums F, Mensah N, Mpembeni R, Kaltschmidt J, Haefeli WE, Blank A. Health workers' knowledge of and attitudes towards computer applications in rural African health facilities. *Glob Health Action* 2014;7:24534 [FREE Full text] [Medline: [25361721](https://pubmed.ncbi.nlm.nih.gov/25361721/)]
38. Wolff-Piggott B, Coleman J, Rivett U. The clinic-level perspective on mHealth implementation: a South African case study. *Information Technology for Development* 2017 Jan 06;24(3):532-553 [FREE Full text] [doi: [10.1080/02681102.2016.1233858](https://doi.org/10.1080/02681102.2016.1233858)]
39. Malema RN, Malaka DW, Mothiba TM. Experiences of lay counsellors who provide VCT for PMTCT of HIV and AIDS in the Capricorn District, Limpopo Province. *Curatiosis* 2010 Sep;33(3):15-23. [doi: [10.4102/curatiosis.v33i3.3](https://doi.org/10.4102/curatiosis.v33i3.3)] [Medline: [21428235](https://pubmed.ncbi.nlm.nih.gov/21428235/)]
40. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. *Res Nurs Health* 2010 Apr;33(2):164-173. [doi: [10.1002/nur.20373](https://doi.org/10.1002/nur.20373)] [Medline: [20198637](https://pubmed.ncbi.nlm.nih.gov/20198637/)]

## Abbreviations

- ART:** antiretroviral therapy
- CFIR:** Consolidated Framework for Implementation Research
- FG:** focus group
- LC:** lay counselor
- mHealth:** mobile health
- PDC:** People Development Center
- TB:** tuberculosis

*Edited by L Buis; submitted 12.08.20; peer-reviewed by R Keith, L Marc, LS Tso; comments to author 17.11.20; revised version received 04.01.21; accepted 23.02.21; published 06.04.21.*

*Please cite as:*

*McCreesh-Toselli S, Torline J, Gouse H, Robbins RN, Mellins CA, Remien RH, Rowe J, Peton N, Rabie S, Joska JA  
Staff Perceptions of Preimplementation Barriers and Facilitators to a Mobile Health Antiretroviral Therapy Adherence Counseling  
Intervention in South Africa: Qualitative Study*

*JMIR Mhealth Uhealth 2021;9(4):e23280*

*URL: <https://mhealth.jmir.org/2021/4/e23280>*

*doi: [10.2196/23280](https://doi.org/10.2196/23280)*

*PMID: [33821806](https://pubmed.ncbi.nlm.nih.gov/33821806/)*

©Siobhan McCreesh-Toselli, John Torline, Hetta Gouse, Reuben N Robbins, Claude A Mellins, Robert H Remien, Jessica Rowe, Neshaan Peton, Stephan Rabie, John A Joska. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 06.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Influences on the Uptake of Health and Well-being Apps and Curated App Portals: Think-Aloud and Interview Study

Dorothy Szinay<sup>1</sup>, MSc; Olga Perski<sup>2</sup>, PhD; Andy Jones<sup>1</sup>, PhD; Tim Chadborn<sup>3</sup>, PhD; Jamie Brown<sup>2,4</sup>, PhD; Felix Naughton<sup>1</sup>, PhD

<sup>1</sup>School of Health Sciences, University of East Anglia, Norwich, United Kingdom

<sup>2</sup>Department of Behavioural Science and Health, University College London, London, United Kingdom

<sup>3</sup>Behavioural Insights, Public Health England, London, United Kingdom

<sup>4</sup>SPECTRUM Consortium, London, United Kingdom

**Corresponding Author:**

Dorothy Szinay, MSc  
School of Health Sciences  
University of East Anglia  
Norwich  
United Kingdom  
Phone: 44 1603593064  
Email: [d.szinay@uea.ac.uk](mailto:d.szinay@uea.ac.uk)

## Abstract

**Background:** Health and well-being smartphone apps can provide a cost-effective solution to addressing unhealthy behaviors. The selection of these apps tends to occur in commercial app stores, where thousands of health apps are available. Their uptake is often influenced by popularity indicators. However, these indicators are not necessarily associated with app effectiveness or evidence-based content. Alternative routes to app selection are increasingly available, such as via curated app portals, but little is known about people's experiences of them.

**Objective:** The aim of this study is to explore how people select health apps on the internet and their views on curated app portals.

**Methods:** A total of 18 UK-based adults were recruited through social media and asked during an in-person meeting to verbalize their thoughts while searching for a health or well-being app on the internet on a platform of their choice. The search was then repeated on 2 curated health app portals: the *National Health Service Apps Library* and the Public Health England *One You* App portal. This was followed by semistructured interviews. Data were analyzed using framework analysis, informed by the Capability, Opportunity, Motivation-Behavior model and the Theoretical Domains Framework.

**Results:** Searching for health and well-being apps on the internet was described as a *minefield*. App uptake appeared to be influenced by participants' capabilities such as app literacy skills and health and app awareness, and opportunities including the availability of apps, app esthetics, the price of an app, and social influences. Motivation factors that seemed to affect the uptake were perceived competence, time efficiency, perceived utility and accuracy of an app, transparency about data protection, commitment and social identity, and a wide range of emotions. Social influences and the perceived utility of an app were highlighted as particularly important. Participants were not previously aware of curated portals but found the concept appealing. Curated health app portals appeared to engender trust and alleviate data protection concerns. Although apps listed on these were perceived as more trustworthy, their presentation was considered disappointing. This disappointment seemed to stem from the functionality of the portals, lack of user guidance, and lack of tailored content to an individual's needs.

**Conclusions:** The uptake of health and well-being apps appears to be primarily affected by social influences and the perceived utility of an app. App uptake via curated health app portals perceived as credible may mitigate concerns related to data protection and accuracy, but their implementation must better meet user needs and expectations.

(*JMIR Mhealth Uhealth* 2021;9(4):e27173) doi:[10.2196/27173](https://doi.org/10.2196/27173)

**KEYWORDS**

behavior change; health apps; mHealth; smartphone app; framework analysis; Capability, Opportunity, Motivation-Behavior model; Theoretical Domains Framework; think aloud; mobile phone

## Introduction

### Background

Noncommunicable diseases (eg, diabetes, heart disease and cancer as well as poor mental health) are considered key threats to global health [1] and are driven by factors such as physical inactivity, poor diet, tobacco smoking, and excessive alcohol consumption. A key global public health policy priority is to enact policies to ensure that the best possible health care is available for all [2]. In the United Kingdom, aims of the National Health Service (NHS) long-term plan [3] and priorities of UK government executive agencies such as Public Health England (PHE) are to provide a smoke-free society, to encourage healthier diets, and to improve mental health [4]. Encouraging the use of digital health interventions, such as smartphone apps, may be a cost-effective way of contributing.

Health and well-being smartphone apps can be cost-effective solutions for changing health behaviors [5,6]. Such tools can act as ideal platforms to deliver behavior change interventions [7] because of their availability, portability, and easy access [8]. Research has demonstrated early evidence of effectiveness of smartphone apps for smoking cessation [9], healthy dietary and physical activity promotion [5,10-12], weight loss [5,13,14], alcohol reduction among nondependent drinkers [15], and mental health promotion [16]. In addition, health apps can reach those resistant to seeking help in person (because of stigma) by improving access to behavior change interventions [17]. However, low uptake and poor engagement over time compromise the potential of health and well-being apps.

*Uptake* refers to the decision to select and install a health app [18]. The search for and selection of health apps tend to take place in commercial app stores such as Google Play for Android operating systems and the Apple App Store for iOS [10,19]. Thousands of health and well-being smartphone apps are available in the major app stores, a number that continues to grow [7], and the uptake of apps from commercial app stores tends to be influenced by indicators of popularity such as the app's rank order, ratings and reviews, and the total number of downloads [19]. However, such popularity indicators are not necessarily positively associated with the effectiveness of an app [20] and may even be negatively related [21]. An associated problem with app uptake is that the vast majority of apps listed in commercial stores lack evidence about their efficacy [22] or effectiveness [23]. The need for quality marks in commercial app stores [24] and regulation of health apps and evidence for their effectiveness has been raised [16]. Better transparency in an app's description to help people make an informed choice,

including how the user's data are handled, how the app was developed, benefits explained in lay terms, and descriptions of the app content, has been recommended [25-27].

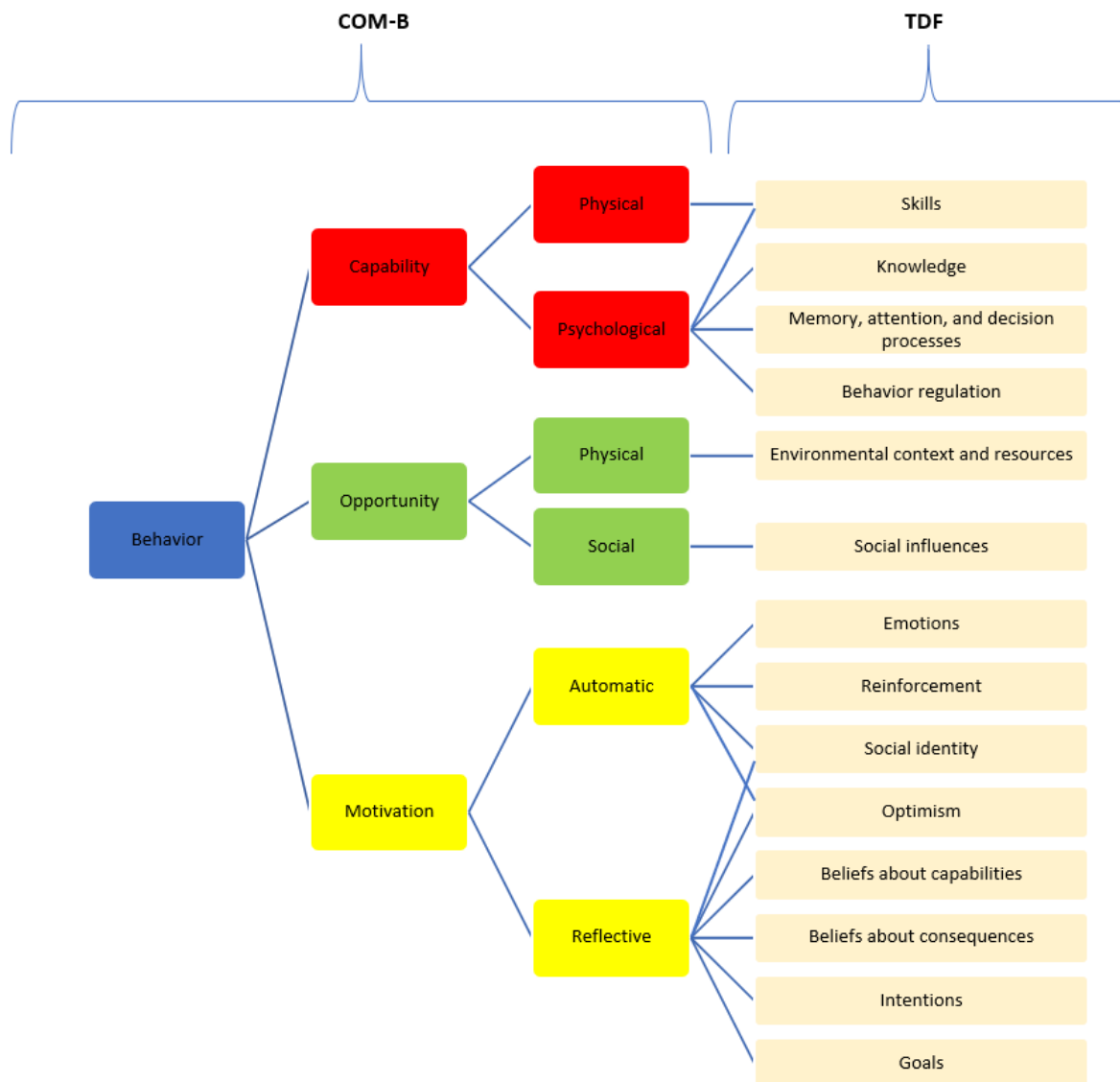
A barrier to the uptake of evidence-informed apps is that not all apps are available to the public, or prominently displayed, via commercial app stores [22,24]. Therefore, fewer people may benefit from the high-quality tools available. Evidence-informed apps tend to be promoted within community or health care settings (often targeting a specific geographic region or country) or on curated health app *portals*. These portals are websites that present a list of selected health apps [28]. Health app portals can be government funded, such as the UK NHS's *Apps Library* or PHE's *One You Apps* portal, or curated by private organizations, such as *App Script* by IQVIA in the United States, the United Kingdom, and the United Arab Emirates; the *MyHealthApps* by PatientView's in Europe and the United Kingdom; or *ORCHA Health* in the United Kingdom. These organizations can lend credibility to and have the potential to promote the uptake of selected health apps [29] by providing a list of safe; evidence-informed; tested; and, where possible, clinically effective health apps for the general public to choose from.

Research has focused on the identification of factors that influence the uptake of health apps in commercial app stores. There is an urgent need to explore whether the general public would be willing to use curated health app portals, which could improve the uptake of evidence-informed health and well-being apps [18]. Despite this need, little is known about the views on curated health app portals. This study aims to explore potential users' views on factors influencing the uptake of health apps in general and on curated health app portals in particular using think-aloud and interview methodology.

### Theoretical Framework

The Capability, Opportunity, Motivation-Behavior (COM-B) model [30] offers a comprehensive framework for understanding behaviors. In the context of this study, the behavior of interest is the uptake of health and well-being apps. The model proposes that behavior arises because of the interaction of three components: capability (physical and psychological), opportunity (physical and social), and motivation (automatic and reflective). The Theoretical Domains Framework (TDF) [31], which contains 14 domains that can be mapped onto the components of the COM-B model, was also used. Together, the COM-B model and the TDF allow for a detailed analysis of data and identification of key factors influencing uptake in general and on curated health app portals in particular (Figure 1) [18].

**Figure 1.** A visual representation of mapping the TDF onto the COM-B model. COM-B model: Capability, Opportunity, Motivation-Behavior model; TDF: Theoretical Domains Framework.



**Aims**

This qualitative study applied a theoretical framework informed by the COM-B model and TDF to explore (1) factors influencing potential users’ uptake of health and well-being smartphone apps through searching on the internet and (2) their views on available curated health app portals.

**Methods**

**Study Design**

This study elicited views and preferences of a sample of members of the public. The Consolidated Criteria for Reporting Qualitative Research checklist guided the design of the study [32] (checklist given in Multimedia Appendix 1). The think-aloud methodology [33] was applied to collect real-time data about health app selection on the internet and involved asking participants to verbalize their thoughts and impressions throughout the selection process. The researcher intervened only when a prompt was considered necessary (eg, during silent

moments, asking questions such as “What are you thinking now?”). Following the think-aloud tasks, follow-up questions were asked to better understand the statements or utterances made during the tasks. Finally, semistructured interviews were conducted. The think-aloud tasks and the topic guide were informed by stakeholder consultation, which included views and opinions of lay persons (patient and public involvement representatives) and expert opinions of policy makers of this study. The study protocol was preregistered on the Open Science Framework [34]. The Faculty of Medicine and Health Sciences Ethics Committee at the University of East Anglia approved this study (reference number: 201819-089). The collected data are stored following the European Union General Data Protection Regulation and the University of East Anglia Research Data Management Policy. The data were anonymized, and all personal identifiers were removed. All participants read the participant information sheet and provided consent before participating in the study.

## Participants and Recruitment

Participants were recruited through paid advertisements on Facebook. Adults in the general population were eligible if they were 18 years or older; were able to provide consent; owned a smartphone; would consider using a smartphone app to change their behavior in the future; and were able to attend an interview in Norwich, England, where the work took place. As a standard practice in qualitative research, the aim of this study is to gain a better understanding of the phenomenon of interest and to increase the coverage of perspectives rather than to recruit a population-representative sample [35]. Therefore, purposive sampling was used to promote the diversity of the sample (ie, age, gender, ethnicity, educational level, and employment) [36]. This included targeted advertisements on Facebook and the selection of participants to ensure the diversity of the sample. A total of 114 individuals responded to the Facebook advertisements and read a brief participant information sheet and completed the screening questionnaire. Of the 38 participants invited to an interview, 14 did not respond and 24 agreed to participate. Of these 24 participants, 6 were canceled for various reasons.

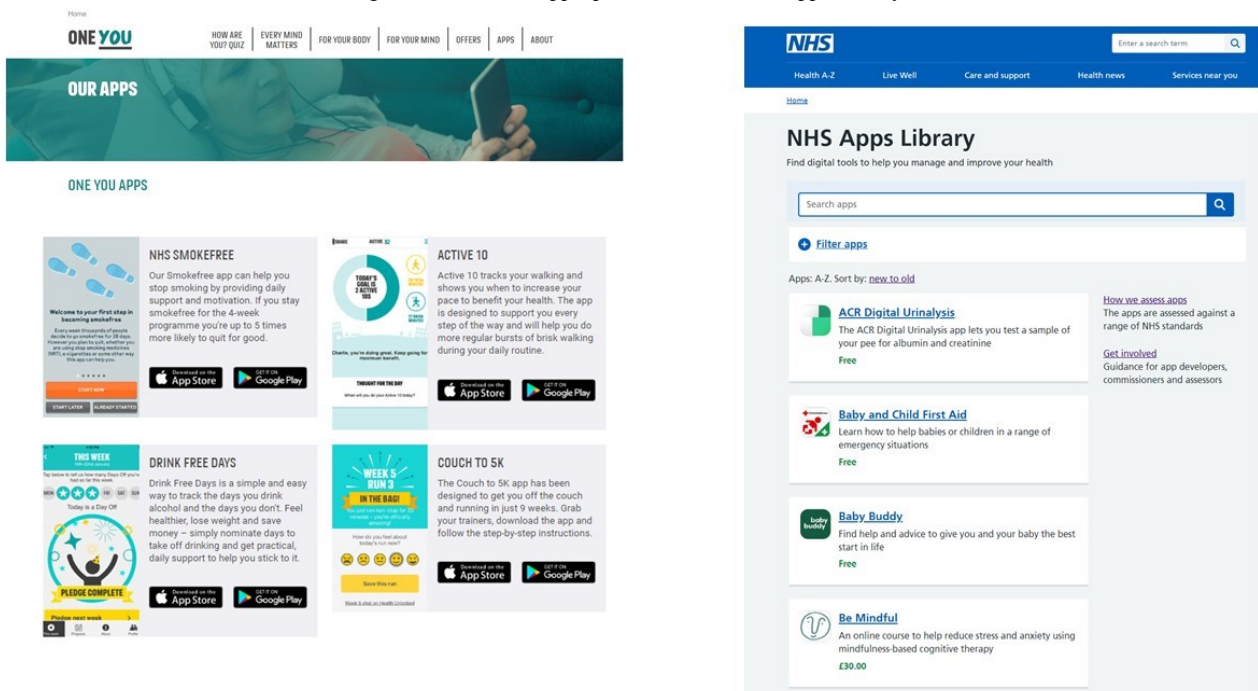
## Procedure

Before completing the online screening survey, participants were asked to read a brief participant information sheet describing the study. After reading and agreeing to participate, participants were asked to complete an online questionnaire to assess their eligibility and to collect descriptive data (Multimedia Appendix 2). Data were collected on age; gender; ethnicity, measured using the Office for National Statistics' index; level of education; employment status; whether they had ever used health or well-being apps; whether they currently use a health

or well-being app; last time they had downloaded an app; and frequency of app use. Participants who met the inclusion criteria were sent an email with a comprehensive participant information sheet and invited to participate in the interview. On the day of the interview, the interviewees received a printed copy of the participant information sheet, and written consent was obtained.

Face-to-face interviews were conducted between July and August 2019 and took place at the University of East Anglia (n=17) or participants' homes in Norwich (n=1). The interviews were conducted by a single female researcher (DS), and no one else was present during the sessions. Each session started with a think-aloud exercise, with participants being instructed on how to verbalize their thoughts. First, they were asked to perform a search for an app they would potentially use to change the health behavior of their choice. They had a choice of using either a study laptop or their smartphone. Second, the researcher asked them if they were familiar with curated app portals. If they were not, DS briefly explained the principle and asked them to repeat the search using the *NHS Apps Library* and the PHE's *One You Apps* curated health app portals (Figure 2). During the think-aloud sessions, positive reinforcement using verbal (eg, "You are doing great" and "Right") and nonverbal (eg, nodding) communication was used to encourage participants to continue to express their views. In quiet moments, prompts were used (eg, "What are you thinking now?" and "Tell me what is on your mind"). Following the think-aloud task, questions regarding their experience with the uptake of and engagement with apps were asked (the topic guide is given in Multimedia Appendix 3). The sessions lasted between 26 and 63 minutes. Participants received a US \$27.50 (UK £20) gift voucher as compensation for their time.

Figure 2. Screenshot of the Public Health England's 'One You Apps' portal and the 'NHS Apps Library'.



## Data Analysis

The sessions were audio recorded and transcribed verbatim by an external company. The transcriptions were checked for accuracy by the researcher undertaking the interviews. The data were analyzed using framework analysis following the stages of familiarization, identification of thematic framework, indexing, charting, mapping, and interpretation [37]. To ensure rigor, trustworthiness, and consistency, a percentage of randomly selected transcripts (2/18, 15%) were independently coded by the second author (OP). The deductive thematic framework based on TDF was refined iteratively through repeated discussions with the second author (OP), and any discrepancies were resolved through discussion with the senior author (FN). Indexing was completed by the first author (DS) using QSR NVivo 12. The data were charted, and the responses were grouped according to the finalized thematic framework. During mapping and interpretation, the grouped data were examined by DS to identify patterns. During mapping, identified factors were classified according to their organic position rather than what they affect (eg, an opportunity factor may indirectly influence the behavior by increasing the motivation for uptake of a health app and influencing it directly). To aid comprehension of the findings for uptake in general and on health app portals in particular, data were analyzed and presented separately for these 2 topics.

## External Validity

To enhance the credibility and trustworthiness of the results [38], 30% (6/18) of participants were randomly selected and requested via email to provide feedback on a document with a summary of the findings and conclusions (*member checking*). They were asked whether they recognized their opinions and whether they agreed with the interpretation of the findings. A total of 2 participants responded to our request and confirmed that their opinions had been captured. In one case, our email was not delivered.

## Reflexivity

The researchers involved in this study are mixed methods researchers with experience applying the COM-B model and TDF to qualitative data. She disclosed her research interest to

participants on the day of the interview, and no previous relationship was established between her and participants. The interviews were conducted by the lead author, a PhD candidate who has undertaken extensive training in the collection and analysis of qualitative data. Participants were encouraged to share their thoughts (both positive and negative) and to be honest. The interviewer felt that good rapport was built with the interviewees, and most participants (n=16) expressed their interest in learning more about the findings of the research. Field notes and a research journal were kept during data collection.

## Results

### Participant Characteristics

A total of 18 participants completed the interview. The average age of participants was 43 (SD 14) years, 50% (n=9) were female, 78% (n=14) were of White British ethnicity, 72% (n=13) were employed full time, 11% (n=2) had postgraduate qualifications, 94% (n=17) had used health apps before, and 61% (n=11) were using health apps at the time of the interviews, out of which 73% (n=8) reported daily health app use. Most participants were interested in changing more than one behavior (eg, losing weight, getting more active, and managing mood), and only 16% (n=2) of participants expressed a desire to change only one behavior. Participants' characteristics are presented in [Multimedia Appendix 4](#).

A total of 2 participants were satisfied with the app they were already using and did not wish to take part in the think-aloud exercise to look for a different app. The remaining 16 participants searched for apps targeting physical activity (n=6), weight management (n=4), mood and mental well-being (n=3), smoking cessation (n=1), alcohol reduction (n=1), and sleep (n=1).

The findings pertaining to factors relevant for both the uptake of health apps and views on curated health app portals are presented under the components of the COM-B model. Higher order themes and subthemes informed by the COM-B model and TDF are reported in [Table 1](#).

**Table 1.** Factors influencing uptake of health apps in general and on health app portals mapped onto the components of the Capability, Opportunity, Motivation-Behavior model and Theoretical Domains Framework constructs.

COM-B <sup>a</sup> component and TDF <sup>b</sup> construct and identified factor	Uptake in general	Uptake on health app portals
<b>Physical capability</b>		
<b>Skills</b>		
App literacy	<ul style="list-style-type: none"> <li>Technological competency</li> </ul>	— <sup>c</sup>
<b>Psychological capability</b>		
<b>Knowledge</b>		
Health awareness	<ul style="list-style-type: none"> <li>General health consciousness or having family members diagnosed with a condition or disease or concerns regarding a behavior or health outcome</li> </ul>	—
App awareness	<ul style="list-style-type: none"> <li>Knowledge of the existence of health and well-being apps</li> </ul>	<ul style="list-style-type: none"> <li>Knowledge of the existence of health and well-being apps listed on health app portals</li> </ul>
User guidance	—	<ul style="list-style-type: none"> <li>Instructions on how to effectively use a health app portal</li> </ul>
Health information	—	<ul style="list-style-type: none"> <li>Educational information related to health and well-being</li> </ul>
<b>Memory, attention, and decision processes</b>		
Cognitive load	—	<ul style="list-style-type: none"> <li>The manner in which apps are presented on the portal</li> <li>The complexity of the search or to access a relevant health app</li> </ul>
<b>Physical opportunity</b>		
<b>Environmental resources</b>		
Availability	<ul style="list-style-type: none"> <li>The ability to use a smartphone anytime, anywhere</li> <li>Availability of an app on all major commercial app stores</li> </ul>	—
Portal tailored to individuals' needs	—	<ul style="list-style-type: none"> <li>Personalized listing of apps targeting age, gender, and health condition</li> </ul>
Cost of an app	<ul style="list-style-type: none"> <li>Low cost and apps that are free for users</li> </ul>	<ul style="list-style-type: none"> <li>Low cost and apps that are free for users</li> </ul>
Esthetics	<ul style="list-style-type: none"> <li>The look and design of an app</li> </ul>	<ul style="list-style-type: none"> <li>User-friendly and design-related characteristics of the portal</li> </ul>
<b>Social opportunity</b>		
<b>Social influences</b>		
Social influences	<ul style="list-style-type: none"> <li>The importance of reviews and ratings in the commercial app stores and apps promoted as "editor's choice"</li> <li>Identified credible sources: apps developed or endorsed by trusted app developers, organizations, or universities or promoted by respected celebrities (eg, athletes)</li> <li>Recommendations received from health practitioners or from friends and family</li> </ul>	<ul style="list-style-type: none"> <li>Health app portals perceived as credible sources</li> <li>Recommendations of health app portals needed mainly in primary care</li> <li>Clarity about the recommended apps on health app portals</li> <li>Explanations about any required GP<sup>d</sup> referral</li> </ul>
<b>Reflective motivation</b>		
<b>Beliefs about capabilities</b>		
Perceived competence	<ul style="list-style-type: none"> <li>Apps preferred over face-to-face intervention when the user feels that they can engage with the app on their own</li> </ul>	—

COM-B <sup>a</sup> component and TDF <sup>b</sup> construct and identified factor	Uptake in general	Uptake on health app portals
<b>Beliefs about consequences</b>		
Time efficiency	<ul style="list-style-type: none"> <li>The ability of a health app to be interacted with a minimum amount of time</li> </ul>	—
The perceived utility of the app	<ul style="list-style-type: none"> <li>Discrepancies between what users are looking for and what the app offers, characterized by a relevant title, description, pictures, adaptation to individual characteristics, and users' previous experience with health apps</li> </ul>	<ul style="list-style-type: none"> <li>Discrepancies between what users are looking for and what the app listed on health app portal offers, characterized by a relevant title, description, and pictures</li> </ul>
Perceived accuracy	<ul style="list-style-type: none"> <li>The perceived effectiveness of apps before the selection of an app</li> </ul>	<ul style="list-style-type: none"> <li>Potential app users' perceived effectiveness of apps listed on health app portals</li> </ul>
Data protection	<ul style="list-style-type: none"> <li>Concerns regarding the handling of personal data</li> </ul>	<ul style="list-style-type: none"> <li>Concerns over the handling of personal data</li> </ul>
<b>Intentions</b>		
Commitment	<ul style="list-style-type: none"> <li>The level of commitment when deciding to download a health app</li> </ul>	—
<b>Social identity</b>		
Social identity	<ul style="list-style-type: none"> <li>Identity related to app use (eg, trends and gender specificity)</li> </ul>	<ul style="list-style-type: none"> <li>Identity related to app use (eg, feeling like a "patient")</li> </ul>
<b>Automatic motivation</b>		
<b>Emotions</b>		
Positive	<ul style="list-style-type: none"> <li>Triggered by curiosity in trying a health app, and by the time efficiency characteristic of an app as opposed to face-to-face interventions, and being provided by a credible source</li> </ul>	<ul style="list-style-type: none"> <li>Triggered by curiosity in choosing a behavior change tool from a curated health app portal and from a credible source</li> </ul>
Negative	<ul style="list-style-type: none"> <li>Triggered by lack of availability on all major app stores</li> <li>Preferred over a face-to-face intervention if feeling anxiety (eg, caused by an unhealthy behavior or unhealthy state) and pressurized (to succeed or show progress)</li> </ul>	<ul style="list-style-type: none"> <li>Triggered by lack of search features on the portal or when the search yields irrelevant results; when an app requires GP referral without further explanation or when an app is only available in one major app store</li> </ul>
Mixed	<ul style="list-style-type: none"> <li>Triggered by the esthetics (design) of the apps and by adaptation to individual characteristics (judged by the title, description, pictures, and gender specificity)</li> </ul>	<ul style="list-style-type: none"> <li>Triggered by the esthetics and features of the portal and the perceived utility of the apps</li> </ul>

<sup>a</sup>COM-B: Capability, Opportunity, Motivation-Behavior.

<sup>b</sup>TDF: Theoretical Domains Framework.

<sup>c</sup>Not available.

<sup>d</sup>GP: general practitioner.

## Factors Influencing the Uptake of Health and Well-being Apps

Half of the participants who agreed to search for a health app (n=8) used Google Search as their first choice to find a suitable app, whereas the other half opened a commercial app store. The latter search among hundreds of available apps was described by most participants as difficult or a "minefield" (P2, P4, and P6). One participant described this task as being "far more complicated than I thought it would be" (P2). By the end of this exercise, only 3 participants found an app that they were

willing to download and engage with further to change their behavior.

### *Capability Factors Related to the Uptake of Health and Well-being Apps in General*

Participants who presented a higher level of technological competency were able to better navigate on their phones, thus highlighting that app literacy skills are necessary when selecting a health app. One participant, who had never used a health app before, showed signs of technical difficulties (ie, lack of skills) during the think-aloud exercise while searching for an alcohol reduction app in a commercial app store:

*I wouldn't know how to do that [refining the search to find a suitable app]. [P12]*

In addition, 2 participants expressed their concern toward the older generation and stated that training should be provided for those with insufficient technological and app literacy skills:

*My nanny is diabetic and if there was an app to help her with her diabetes, then I'm sure she would be happy to use it but it's just someone would need to explain it to her. [P18]*

All participants expressed their decision to look for an app for health reasons, such as getting healthier or preventing illness. This included reasons of being diagnosed, or having a family member diagnosed, with a medical condition (eg, diabetes and high blood pressure) or concerns of the negative effect a current behavior may have (eg, smoking and alcohol consumption) to better manage or improve their mental health (eg, anxiety and self-confidence) and general well-being (eg, sleep quality):

*I'm trying to avoid having type 2 diabetes, or getting it, so there's a background, my mother, in my family, there's a heart conditions background, which is why I'm really wanting to do something about my health. [P3]*

Although most participants were aware of the existence of some apps, 3 participants were surprised by the existence of health apps for smoking cessation and mental health issues:

*It didn't cross my mind that I could use an app for stopping smoking, so it is new. [P16]*

### **Opportunity Factors Related to the Uptake of Health and Well-being Apps in General**

Some participants expressed their preference to look for a health app as a digital behavior change intervention instead of a face-to-face intervention because of the availability and low cost of an app. However, concerns around widening inequalities were raised by one participant who showed signs of worry about the limited access to digital aids for individuals living in deprived areas:

*So if they [people living in deprived areas] do not have the smart phone, they won't be able to use it, so it's not going to work, is it? It's what happened with the Universal Credit, so it's not going to work. I mean issue everyone a smart phone. [P16]*

A few participants highlighted the importance of the availability of health apps in both major commercial app stores (Apple App Store and Google Play), not just one or the other.

Most participants stated that apps should be available at no cost. Only 6 participants expressed their willingness to pay a small fee for an app if, for example, it would be “almost life-changing” (P4) or if it would include online professional support.

The specific design and color scheme preferred by participants appeared to be unique and dependent on the individual's taste. However, the majority were looking for a *simple* looking app.

Social influences appeared to be one of the core factors that shaped the selection of apps for all participants during the think-aloud exercise. This includes ratings and reviews of the

app, the credibility of the source of the app, and recommendations of apps received from others. Within app stores, most participants described looking at the star ratings and the number of downloads of each app and whether the apps were listed as an *editor's choice*. A total of 3 participants acknowledged that reviews were subjective, and they still reported feeling influenced by the ratings of the app. In addition, 2 participants reported that they were skeptical of the reviews, which they believed may have been paid for, and that reviews are not enough, as more information is necessary to make an informed choice:

*You know, so you're having to make all these judgements about people's reviews and then you know deep down that the reviews might be paid for and, you know, it's a bit of a minefield which is why I would only take a free sample and then see if it works for me. [P6]*

A credible source was also important. Apps developed or recommended by trusted organizations or respected celebrities seemed more appealing to all participants. Participants who used Google Search to find an app aimed to look for websites they were familiar with or had used before or for websites that would post “Top 10 apps for...” type of articles. In addition, word of mouth was another source of social influence:

*I see two different specialists, I have a lung problem as well and I see a lung specialist at a hospital near me and she said to me, the best thing that I could do, which was downloading the Couch to 5k app. [P14]*

### **Motivational Factors Related to the Uptake of Health and Well-being Apps in General**

Health or well-being apps were preferred over face-to-face options because participants reported feeling competent by changing their behavior through the use of an app, requiring less time commitment and avoiding the anxiety and pressure of interacting with others. Time appeared to be a particularly valuable resource for all participants, and they believed apps to have this advantage.

Another core factor in the selection of an app was the way users perceived its utility. This was based on 2 aspects. First, they appeared to judge how the app is adapted to the individual by reading the title and description of the app and by looking at pictures (ie, screenshots). A total of 12 participants reported the need for sufficient information about an app to make an informed choice:

*I would definitely judge more from the pictures more than anything and I think that just nowadays everyone does, is you get an idea of the app from the pictures. (...) I mean I think when you see an older person on a picture and you're a lot younger, it makes you think, I mean it's the wrong think to think but it makes you think maybe it's not for me. [P7]*

Second, it seemed that 12 participants relied on their past experiences with health apps. Whether those experiences were positive or negative may have shaped their beliefs about health apps in general:



*So that's why My Fitness Pal is the first app that I've ever had that's actually worked. [P9]*

In addition, 7 participants expressed skepticism about the accuracy and effectiveness of some apps (eg, mental health apps), and concerns about data protection were mixed:

*These mindful ones, I've never downloaded one and I'm sceptical. [P17]*

Participants mentioned that commitment to the behavior change would influence uptake and future engagement:

*So I think the committed ones seek out the ones that are the right ones for them, the best ones, rather than necessarily the trendy ones. [P4]*

Participants' social identities also shaped their selections. Many reported that they did not wish to select apps that promoted groups they did not seem to fit in with (eg, athletic body image or individuals of the hipster subculture):

*They've got a kind of hipster bloke and now they've got a kind of sexy female image with tattoos down her arm, sexy, trendy, female image. Okay, so they are obviously aiming at younger, sort of people in their twenties and thirties, yeah, another sexy female image. It's quite interesting isn't it, I'm looking at the images and not the words and getting a sense, is this for me, middle aged, well older woman?! [P6]*

Curiosity, defined here as a desire to learn something, was the only stand-alone positive emotion and appeared to positively influence the uptake of health apps for many participants:

*I thought out of curiosity I'd have a look, so I just typed in quit smoking in Google play store and there's hundreds of apps from various people with varying degrees of credibility, and they all were pretty similar to be honest. [P13]*

Apps linked to a credible source were important, with people unimpressed when an app was not available on all major app stores.

### Views on Curated Health App Portals

None of the participants spontaneously used a curated portal. Curated portals were then introduced to the participants, but none were previously aware of them. Curated health app portals were appealing to all participants, and they believed the portals would be likely to engender trust. However, searching for a health app on the NHS Apps Library and the One You App portal was a generally disappointing experience. Only 2 participants chose a health app from a health app portal (One You Apps), whereas the rest of the participants decided to continue the search in commercial app stores.

### Capability Factors Related to the Uptake of Health and Well-being Apps on Health Portals

All participants had heard of widely advertised apps (eg, Couch to 5k), but none were aware of the existence of curated health app portals before participating in this study:

*I think they're brilliant [apps on health app portals]; I didn't know they existed. [P11]*

Navigating on the NHS Apps Library seemed easy for some. However, a few participants mentioned that a user guide or help section would be a useful added feature of the portal. Two participants reported that they did not find it easy to use the filter features, and in many cases, they felt the search yielded irrelevant results (eg, while searching for a physical activity app, the results also listed apps for mental health). A few participants reported that navigating on curated app portals was difficult, characterized as "cumbersome" (P4, P12):

*It's not clear, it's suggests that they are independent apps but maybe they should have some guidelines about design, you know, of their sort of landing pages. [P6]*

### Opportunity Factors Related to Uptake of Health and Well-being Apps on Health Portals

All participants indicated that they would want a portal tailored to their needs, with categories related to their gender, age group, and medical conditions they may have:

*So something like that, this is suitable if you're over 65, this would be more suitable for you if you're under 40 or with these ones that you don't have to go and see your GP, that you can pay for, if you have any concerns, visit your GP or speak to a health professional because some people don't have that common sense. [P14]*

Participants had different opinions about the layout of these portals. Some liked the NHS Apps Library design better, with simple colors, whereas others enjoyed the more colorful One You App portal. Most participants felt that a fusion between these 2 designs (the searchability and filters of the NHS Apps Library and the look and presentation of the One You App portal) and a better functionality would create the ideal curated health app portal:

*Why they are not combined? [P8]*

Although many participants expressed their wish to access apps for free, a few participants were more open to pay for an app that was listed on a curated health app portal:

*This is fabulous, and I'd be much more inclined to pay money. This is really, really good. [P6]*

Participants found the NHS and PHE trustworthy and believed that these portals would provide safe and effective digital aids. Some indicated a desire to receive further recommendations for using these portals from their primary care physicians:

*If GPs knew that they could say "well this could help you" I'm sure that they would recommend it to people. [P11]*

However, they also wanted to avoid putting unnecessary pressure on general practitioner (GP) practices:

*You've got "free but requires GP referral" and when you're thinking the NHS is under so much financial strain and pressure at the moment, why do I need a GP referral to obtain an app? [P2]*

In addition, the One You App portal lists a few apps that are recommended, but participants expressed their confusion and

lack of clarity regarding why some apps are *recommended* and by whom.

### **Motivation Factors Related to Uptake of Health and Well-being Apps on Health Portals**

While searching on curated health app portals, none of the participants expressed signs of concern about data protection and accuracy of apps, although 2 participants reported that they would want to read more about how these apps were developed and tested:

*How long it takes, how many sessions and the fact that it's been tested in clinical trials and evaluated by NICE which, to me, is probably quite an important thing. [P1]*

Social identity was also important. Some participants had identified themselves as individuals living with a medical condition. These participants were keen to look for an app that targets the behavioral change of individuals with preexisting medical conditions. Others stated that they do not wish to feel “like a patient” (P7) and seemed reluctant to continue the search on a curated health app portal:

*So it would be nice to have one specific for maybe people with medical problems or age-related problems, etc. [P15]*

## **Discussion**

### **Principal Findings**

Online searches for health and well-being apps were found to be difficult. Factors influencing the uptake of health apps were mapped using the COM-B model and TDF. We found that social influences and participants' beliefs about consequences (the perceived utility of the app) are key factors influencing the uptake of health apps. This conclusion was based on the frequency and salience of the themes that occurred during the interview. Curated health portals were found to be appealing despite the lack of awareness of their existence. However, the way apps are currently presented on these portals did not meet users' needs because of a lack of certain features, such as lack of tailoring to the user's requirements.

In line with previous research, the findings revealed the importance of the capability and opportunity factors, such as app literacy skills; health awareness and app awareness; esthetics of an app; low cost of an app; reading reviews and checking ratings; credible sources; and recommendations of apps from others, including health professionals [18,22,39,40]. Interestingly, the perception of the cost of an app appeared to be related to the perceived utility and credibility of the source. Although at the start, some participants were against paying for apps, the more useful an app was perceived, the more inclined participants felt to pay a fee. This phenomenon was observed for apps listed on health app portals, which were considered a credible source. More importantly, unlike apps listed on commercial app stores, there was implied trust in apps listed on curated health app portals by participants. In addition, some health apps are not available for downloading in both commercial app stores. Participants found it disappointing that

some apps were only available for iPhone users. This is in line with previous research that found that out of 18 investigated health apps, only one-third were available to download on both major commercial app stores [28].

In terms of motivational factors, we found that perceived utility included aspects related to individuals' perceptions about the presentation of an app and their previous experiences with health apps. Together, these shaped the way participants judged the usefulness of an app. This characterization underlines the need expressed by others previously for a better way to present health apps through a description that would lead to an informed choice (eg, the content of the app) [25-27] and potentially positively affect other motivational factors, such as the accuracy of an app and data protection [41]. Notably, concern about data protection and the accuracy of a health app was minimal when participants navigated on health app portals as opposed to commercial app stores.

There is a need to understand what design aspects generate positive or negative emotions and for whom. Emotions are powerful drivers of a behavior, which affects decision making (eg, app uptake) [42]. A key emotion identified in this study directly influencing the uptake was curiosity. However, this study emphasized the importance of positive emotions triggered by, for example, the credible source of an app and negative emotions triggered by restriction of information (eg, lack of understanding of the necessity of GP referral to download an app). Taking these factors into consideration may lead to better uptake with such tools.

Uptake and engagement are connected. Engagement without uptake is not possible, and uptake without taking into consideration the factors that are important for engagement is impractical. Some factors might influence both uptake and engagement; for example, our research suggests that the perceived utility of an app is one of the main factors for uptake. However, a previous study found that perceived utility was a predictor of engagement with an alcohol reduction app [43]. Therefore, where possible, uptake and engagement should be considered together as 2 linked constructs.

### **Strengths and Limitations**

The main strength of this study lies in its methodology. Given that the aim of this study is to explore uptake with health apps and by applying a user-centered approach, the think-aloud methodology was the appropriate technique to use [33,44] as it will minimize recall bias. Involving stakeholders—patient and public engagement representatives and policy makers—in the design of the research enhances scientific rigor. The purposive sampling technique adopted enabled the recruitment of a wide range of participants that included the same number of females and males and having different levels of education and employment status, and the sample overrepresented ethnicity relative to local rates. The use of the COM-B and TDF to guide the data analysis is another strength of this study.

This study had several limitations. First, asking participants to perform the think-aloud task under observation may not be fully analogous to how they would perform a search when on their own. Second, some identified factors were difficult to define

and describe because of the lack of specificity of the description provided by participants. These include esthetics of apps, often described vaguely (*nice* and *elegant*) and the cognitive load associated with engagement with these (*easy to use*). Third, for a qualitative research study exploring such a broad topic, we felt that information saturation was reached; however, it is possible that additional participants with more varied characteristics would have allowed us to identify additional concepts. Finally, during external validation, a randomly selected subsample of participants was asked via email to provide feedback on the summary of the findings. A total of 50% (3/6) of participants did not reply, and it is unclear whether these participants ignored our request or did not agree with the interpretation of the results.

### Implications for Research, Policy, and Practice

This study has important implications for stakeholders in public health and policy makers who target prevention and health promotion using digital technologies and governmental bodies and trusted health organizations that provide curated health app portals. Low awareness, low app literacy skills, lack of availability on all major app stores, and lack of recommendation in primary care were identified as factors limiting the uptake of health apps in general and on curated app portals. These factors are important for improving the uptake of health apps. Selection was described as difficult. Therefore, there is a need for public guidance on how to identify evidence-based tools [18,22] and

for health practitioners to promote and advise their patients on how to select appropriate health and well-being apps [40]. Raising awareness of such tools through both online and offline promotion channels might provide better access to effective apps.

Our findings could also help developers to reconsider the ways in which apps are currently presented on commercial app stores and app portals, which might, in turn, increase the uptake of evidence-informed health apps. The idea of selecting an app from a health app portal was appealing to all participants, although individuals' needs were not met. These findings describe essential barriers and facilitators related to participants' capability, opportunity, and motivation to take up health and well-being apps. For example, app descriptions and presentations that better align with individuals' needs may increase the uptake of health apps on health app portals. These findings can also be used to inform the development of interventions that specifically aim to promote the uptake of and engagement with evidence-informed health and well-being apps, a priority within the NHS long-term plan (ie, *digital first*). By targeting the identified psychological influences on app uptake through further interventional work, organizations that provide app portals (eg, the NHS and PHE) should be able to increase their impact by helping people to better select appropriate apps. A summary of the recommendations for policy makers, providers, and developers is presented in [Textbox 1](#).

**Textbox 1.** Recommendations for policy makers, industry, health care providers, and app developers based on the Capability, Opportunity, Motivation-Behavior model for a better uptake of health and well-being apps.

#### Capability

- Improve app literacy skills, with a focus on older and marginalized populations, and continue working toward reducing the digital divide (eg, through the use of an outreach approach to target older, migrant, and homeless populations).
- Increase awareness of effective health apps and curated health app portals through promotion online and offline in primary care, mass media, and public spaces.
- Provide guidance on how to use a health app portal (eg, through incorporating an extensive help section) and additional physical and mental health-related evidence-based papers.
- Promote reduced cognitive load on curated health app portals (eg, through the use of images and short app descriptions).

#### Opportunity

- Ensure evidence-informed apps are available for free or at a low cost to everyone.
- Make apps available on all major app stores simultaneously.
- Offer the possibility to tailor the health app portal to target certain demographics (eg, apps for physical activity for women aged 60 years or more).
- Offer apps at low cost and provide explanation for those that require referrals and justifications for the cost of paid apps on curated health app portals.
- Collaborate with interaction design experts and end users to enhance the esthetics of health app portals.
- Promote evidence-informed apps via trusted organizations and provide information on how the apps were developed and tested.
- Encourage health professionals and practitioners of promotion of evidence-informed health apps and health app portals.

#### Motivation

- Provide relevant and realistic titles and avoid general app descriptions. Descriptions should be short but must contain details of what the app offers and how it is able to help the user.
- Provide pictures of the app (eg, screenshots) and avoid pictures that promote an unrealistic body image.
- Provide information about the accuracy and effectiveness of the app (eg, details about development and developers) and how users' data are handled.
- Take into account users' emotions about certain features by constantly involving the users in the development of health apps.

## Future Research

Future research is needed to minimize factors limiting uptake, such as low awareness, low app literacy skills, and a lack of recommendations in primary care. Our results suggest that there is a need to better tailor the design and content of health app portals to better meet individuals' needs. However, the mixed views on specific app designs indicate that more research is needed to investigate whether there are general design principles that are missed and could be followed to accommodate the majority of people or whether better tailoring and/or adaptive interventions should be considered instead. Future research may also want to consider comparing curated health app portals developed by private organizations with those developed by governmental bodies to investigate whether portal design-related features are considered less or more important than credibility and trust in apps listed on them. Experimental research is needed to assess whether there is a trade-off between credibility, social influences, and perceived utility of the apps presented on curated health app portals. Furthermore, with a growing concern around

widening inequalities [45], solutions should be focused on reducing the digital divide and health inequalities that may appear as a result of the financial constraint of owning a smartphone and lack of sufficient app literacy skills.

## Conclusions

Among the factors mapped under capability, opportunity, and motivation components of the COM-B model, social influences and the perceived utility of an app appear to be the core factors influencing uptake in general and on curated health app portals. Curated app portals are considered trustworthy and serve as a credible source for apps; however, there is disappointment with their current implementation of these portals. Uptake of health and well-being apps on health app portals, as opposed to uptake in general, appears to help address people's concerns regarding data protection and the accuracy of apps. Health organizations that develop app portals may consider targeting the factors identified across the COM-B and TDF as part of additional experimental work, as this could help to increase impact through better selection of appropriate health apps.

## Acknowledgments

The authors would like to thank all participants for their contribution to the study. The authors would like to acknowledge the support on the development of a topic guide received from RP and SH (patient and public representatives), the Deputy Director of PHE Digital, and the PHE Strategy and Planning and PHE Strategy and Innovation lead. The authors are grateful to the University College London Tobacco and the Alcohol Research Group for their expert opinions on the data analysis. DS was funded through a PhD studentship, provided jointly by the PHE and the University of East Anglia (R205853HSC). OP received salary support from Cancer Research UK (C1417/A22962).

## Authors' Contributions

DS, FN, AJ, TC, and JB conceptualized the study design. DS wrote the study protocol with contributions from FN, AJ, TC, and JB. All authors commented on the topic guide. DS undertook recruitment of participants, data collection, data analysis, interpretation, and report writing. OP double-coded a proportion of the transcripts. DS, OP, and FN finalized the final thematic framework. DS prepared the manuscript. All authors read, commented on, and contributed to the final manuscript.

## Conflicts of Interest

JB has received unrestricted research funding to study smoking cessation from pharmaceutical companies that manufacture smoking cessation medications. JB, FN, OP, and DS are unpaid members of the scientific committee for the Smoke Free app and have no financial interest in the app.

### Multimedia Appendix 1

Consolidated criteria for reporting qualitative studies: 32-item checklist.

[PDF File (Adobe PDF File), 150 KB - [mhealth\\_v9i4e27173\\_app1.pdf](#)]

### Multimedia Appendix 2

Screening questionnaire.

[PDF File (Adobe PDF File), 45 KB - [mhealth\\_v9i4e27173\\_app2.pdf](#)]

### Multimedia Appendix 3

Topic guide for the interviews.

[PDF File (Adobe PDF File), 137 KB - [mhealth\\_v9i4e27173\\_app3.pdf](#)]

### Multimedia Appendix 4

Study participants' characteristics.

[PDF File (Adobe PDF File), 105 KB - [mhealth\\_v9i4e27173\\_app4.pdf](#)]

## References

1. World Health Organization. Ten Threats to Global Health in 2019. 2019. URL: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019> [accessed 2020-01-02]
2. World Health Organization. World Health Organization Priorities: Health for all. 2020. URL: <https://www.who.int/dg/priorities/health-for-all/en/> [accessed 2020-06-01]
3. The NHS long term plan. National Health Service (NHS). 2019. URL: <https://www.longtermplan.nhs.uk/> [accessed 2020-07-01]
4. Public HE. Executive summary. Public Health England Strategy 2020-25. 2019. URL: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/830105/PHE\\_Strategy\\_2020-25\\_Executive\\_Summary.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/830105/PHE_Strategy_2020-25_Executive_Summary.pdf) [accessed 2020-08-01]
5. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Smith S. A Review of Smartphone Applications for Promoting Physical Activity. *J Community Med* 2016;2(1):21-32 [FREE Full text] [Medline: 27034992]
6. Coughlin SS, Jacobs M, Thind H, Champagne N, Liu B, Golden MS, et al. On the Need for Research-Tested Smartphone Applications for Reducing Exposures to Known or Suspected Breast Carcinogens in Work and Home Environments. *J Environ Health Sci* 2015;1(4):4-4 [FREE Full text] [doi: 10.15436/2378-6841.15.e004] [Medline: 26688831]
7. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. *J Med Internet Res* 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
8. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. *J Biomed Inform* 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
9. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2016;4:11-47. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]

10. Flores MG, 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;17(11):e253 [FREE Full text] [doi: [10.2196/jmir.4836](https://doi.org/10.2196/jmir.4836)] [Medline: [26554314](https://pubmed.ncbi.nlm.nih.gov/26554314/)]
11. Schoeppe S, Alley S, Van LW, 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](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
12. Nour M, Chen J, Allman-Farinelli M. Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: Systematic Review and Meta-Analysis. *J Med Internet Res* 2016 Apr 08;18(4):e58 [FREE Full text] [doi: [10.2196/jmir.5082](https://doi.org/10.2196/jmir.5082)] [Medline: [27059765](https://pubmed.ncbi.nlm.nih.gov/27059765/)]
13. Lee S, Lindquist R. A Review of Technology-Based Interventions to Maintain Weight Loss. *Telemedicine and e-Health* 2015 Mar;21(3):217-232. [doi: [10.1089/tmj.2014.0052](https://doi.org/10.1089/tmj.2014.0052)]
14. Sherifali D, Nerenberg KA, Wilson S, Semeniuk K, Ali MU, Redman LM, et al. The Effectiveness of eHealth Technologies on Weight Management in Pregnant and Postpartum Women: Systematic Review and Meta-Analysis. *J Med Internet Res* 2017 Oct 13;19(10):e337 [FREE Full text] [doi: [10.2196/jmir.8006](https://doi.org/10.2196/jmir.8006)] [Medline: [29030327](https://pubmed.ncbi.nlm.nih.gov/29030327/)]
15. Choo CC, Burton AAD. Mobile Phone Apps for Behavioral Interventions for At-Risk Drinkers in Australia: Literature Review. *JMIR Mhealth Uhealth* 2018 Feb 13;6(2):e18 [FREE Full text] [doi: [10.2196/mhealth.6832](https://doi.org/10.2196/mhealth.6832)] [Medline: [29439946](https://pubmed.ncbi.nlm.nih.gov/29439946/)]
16. Husain I, Spence D. Can healthy people benefit from health apps? *BMJ* 2015 Apr 14;350:h1887. [doi: [10.1136/bmj.h1887](https://doi.org/10.1136/bmj.h1887)] [Medline: [25873345](https://pubmed.ncbi.nlm.nih.gov/25873345/)]
17. 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;2:118 [FREE Full text] [doi: [10.1038/s41746-019-0188-8](https://doi.org/10.1038/s41746-019-0188-8)] [Medline: [31815193](https://pubmed.ncbi.nlm.nih.gov/31815193/)]
18. 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;1-23 [FREE Full text] [doi: [10.2196/17572](https://doi.org/10.2196/17572)] [Medline: [32348255](https://pubmed.ncbi.nlm.nih.gov/32348255/)]
19. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. *JMIR Mhealth Uhealth* 2015;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
20. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. *J Med Internet Res* 2015;17(5):e118 [FREE Full text] [doi: [10.2196/jmir.4060](https://doi.org/10.2196/jmir.4060)] [Medline: [25977135](https://pubmed.ncbi.nlm.nih.gov/25977135/)]
21. Wyatt JC. How can clinicians, specialty societies and others evaluate and improve the quality of apps for patient use? *BMC Med* 2018 Dec 03;16(1):225 [FREE Full text] [doi: [10.1186/s12916-018-1211-7](https://doi.org/10.1186/s12916-018-1211-7)] [Medline: [30501638](https://pubmed.ncbi.nlm.nih.gov/30501638/)]
22. 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;15(11):e247 [FREE Full text] [doi: [10.2196/jmir.2791](https://doi.org/10.2196/jmir.2791)] [Medline: [24240579](https://pubmed.ncbi.nlm.nih.gov/24240579/)]
23. Radovic A, Vona PL, Santostefano AM, Ciaravino S, Miller E, Stein BD. Smartphone Applications for Mental Health. *Cyberpsychol Behav Soc Netw* 2016 Jul;19(7):465-470. [doi: [10.1089/cyber.2015.0619](https://doi.org/10.1089/cyber.2015.0619)] [Medline: [27428034](https://pubmed.ncbi.nlm.nih.gov/27428034/)]
24. Terhorst Y, Rathner E, Baumeister H, Sander L. «Hilfe aus dem App-Store?»: Eine systematische Übersichtsarbeit und Evaluation von Apps zur Anwendung bei Depressionen. *Verhaltenstherapie* 2018 May 8;28(2):101-112. [doi: [10.1159/000481692](https://doi.org/10.1159/000481692)]
25. 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](https://doi.org/10.2196/13375)] [Medline: [30810534](https://pubmed.ncbi.nlm.nih.gov/30810534/)]
26. Wykes T, Schueller S. Why Reviewing Apps Is Not Enough: Transparency for Trust (T4T) Principles of Responsible Health App Marketplaces. *J Med Internet Res* 2019 May 02;21(5):e12390 [FREE Full text] [doi: [10.2196/12390](https://doi.org/10.2196/12390)] [Medline: [31045497](https://pubmed.ncbi.nlm.nih.gov/31045497/)]
27. Charbonneau DH, Hightower S, Katz A, Zhang K, Abrams J, Senft N, et al. Smartphone apps for cancer: A content analysis of the digital health marketplace. *Digit Health* 2020;6:2055207620905413 [FREE Full text] [doi: [10.1177/2055207620905413](https://doi.org/10.1177/2055207620905413)] [Medline: [32110428](https://pubmed.ncbi.nlm.nih.gov/32110428/)]
28. 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](https://doi.org/10.2196/16741)] [Medline: [32012102](https://pubmed.ncbi.nlm.nih.gov/32012102/)]
29. Russell E, Lloyd-Houldey A, Memon A, Yarker J. Factors Influencing Uptake and Use of a New Health Information App for Young People. *Journal of Technology in Human Services* 2019 Jan 28;36(4):222-240. [doi: [10.1080/15228835.2018.1536911](https://doi.org/10.1080/15228835.2018.1536911)]
30. Michie S, Atkins L, West R. *The Behaviour Change Wheel: A Guide to Designing Interventions*. London: Silverback Publishing; 2014.
31. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implement Sci* 2017 Jun 21;12(1):77 [FREE Full text] [doi: [10.1186/s13012-017-0605-9](https://doi.org/10.1186/s13012-017-0605-9)] [Medline: [28637486](https://pubmed.ncbi.nlm.nih.gov/28637486/)]

32. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357 [[FREE Full text](#)] [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
33. Ericsson KA, Simon HA. How to Study Thinking in Everyday Life: Contrasting Think-Aloud Protocols With Descriptions and Explanations of Thinking. *Mind, Culture, and Activity* 1998 Jul;5(3):178-186. [doi: [10.1207/s15327884mca0503\\_3](https://doi.org/10.1207/s15327884mca0503_3)]
34. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. A qualitative study exploring people's perception of factors influencing the uptake and use of health and wellbeing smartphone apps. *Open Science Framework*. 2021 Jan 29. URL: [osf.io/jrkd3](https://osf.io/jrkd3) [accessed 2021-01-01]
35. Pope C, Mays N. Qualitative Methods in Health Research. *Qualitative Research in Health Car* 2006:e2006-e2011. [doi: [10.1002/9780470750841.ch1](https://doi.org/10.1002/9780470750841.ch1)]
36. Benoot C, Hannes K, Bilsen J. The use of purposeful sampling in a qualitative evidence synthesis: A worked example on sexual adjustment to a cancer trajectory. *BMC Med Res Methodol* 2016 Feb 18;16:21 [[FREE Full text](#)] [doi: [10.1186/s12874-016-0114-6](https://doi.org/10.1186/s12874-016-0114-6)] [Medline: [26891718](https://pubmed.ncbi.nlm.nih.gov/26891718/)]
37. Ritchie J, Lewis J, Nicholls C, Ormston R. Qualitative research practice. A guide for social science students and researcher. New York: Sage; 2013:1-456.
38. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member Checking. *Qual Health Res* 2016 Jul 10;26(13):1802-1811. [doi: [10.1177/1049732316654870](https://doi.org/10.1177/1049732316654870)]
39. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Med Inform Decis Mak* 2017 Dec 28;17(1):25 [[FREE Full text](#)] [doi: [10.1186/s12911-017-0422-8](https://doi.org/10.1186/s12911-017-0422-8)] [Medline: [28241759](https://pubmed.ncbi.nlm.nih.gov/28241759/)]
40. Hickey E, McMillan B, Mitchell C. Practitioners should embrace, not ignore, health apps. *BMJ* 2015 May 07;350:h2336. [doi: [10.1136/bmj.h2336](https://doi.org/10.1136/bmj.h2336)] [Medline: [25953643](https://pubmed.ncbi.nlm.nih.gov/25953643/)]
41. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013;15(4):e86 [[FREE Full text](#)] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
42. Lerner JS, Li Y, Valdesolo P, Kassam KS. Emotion and decision making. *Annu Rev Psychol* 2015 Jan 03;66:799-823. [doi: [10.1146/annurev-psych-010213-115043](https://doi.org/10.1146/annurev-psych-010213-115043)] [Medline: [25251484](https://pubmed.ncbi.nlm.nih.gov/25251484/)]
43. Perski O, Naughton F, Garnett C, Blandford A, Beard E, West R, et al. Do Daily Fluctuations in Psychological and App-Related Variables Predict Engagement With an Alcohol Reduction App? A Series of N-Of-1 Studies. *JMIR Mhealth Uhealth* 2019 Oct 02;7(10):e14098 [[FREE Full text](#)] [doi: [10.2196/14098](https://doi.org/10.2196/14098)] [Medline: [31579022](https://pubmed.ncbi.nlm.nih.gov/31579022/)]
44. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* 2015;17(1):e30 [[FREE Full text](#)] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
45. Blank G, Dutton W, Lefkowitz J. Perceived Threats to Privacy Online: The Internet in Britain. URL: <https://oxis.oi.ox.ac.uk/wp-content/uploads/sites/43/2019/09/OxIS-report-2019-final-digital-PDFA.pdf> [accessed 2020-07-01]

## Abbreviations

**COM-B:** Capability, Opportunity, Motivation-Behavior  
**GP:** general practitioner  
**NHS:** National Health Service  
**PHE:** Public Health England  
**TDF:** Theoretical Domains Framework

*Edited by L Buis; submitted 14.01.21; peer-reviewed by A Robinson, S Haug; comments to author 03.03.21; revised version received 15.03.21; accepted 16.03.21; published 27.04.21.*

### *Please cite as:*

Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F  
*Influences on the Uptake of Health and Well-being Apps and Curated App Portals: Think-Aloud and Interview Study*  
*JMIR Mhealth Uhealth* 2021;9(4):e27173  
URL: <https://mhealth.jmir.org/2021/4/e27173>  
doi: [10.2196/27173](https://doi.org/10.2196/27173)  
PMID: [33904827](https://pubmed.ncbi.nlm.nih.gov/33904827/)

©Dorothy Szinay, Olga Perski, Andy Jones, Tim Chadborn, Jamie Brown, Felix Naughton. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 27.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Leprosy Screening Based on Artificial Intelligence: Development of a Cross-Platform App

Márcio Luís Moreira De Souza<sup>1</sup>, MSc; Gabriel Ayres Lopes<sup>2</sup>; Alexandre Castelo Branco<sup>3</sup>, MD; Jessica K Fairley<sup>4</sup>, MD, MSc; Lucia Alves De Oliveira Fraga<sup>1</sup>, PhD

<sup>1</sup>Multicentre Biochemistry and Molecular Biology Program, Federal University of Juiz de Fora, Governador Valadares-MG, Brazil

<sup>2</sup>Fellowship of PROEX Program/UFJF, Federal University of Juiz de Fora, Governador Valadares-MG, Brazil

<sup>3</sup>Reference Center for Endemic Diseases and Special Programs (SMS/GV), Governador Valadares-MG, Brazil

<sup>4</sup>Emory University School of Medicine, Atlanta, GA, United States

**Corresponding Author:**

Lucia Alves De Oliveira Fraga, PhD

Multicentre Biochemistry and Molecular Biology Program

Federal University of Juiz de Fora

R São Paulo 745 Centro

Governador Valadares-MG

Brazil

Phone: 55 33 33011000

Email: [artigoacm@gmail.com](mailto:artigoacm@gmail.com)

## Abstract

**Background:** According to the World Health Organization, achieving targets for control of leprosy by 2030 will require disease elimination and interruption of transmission at the national or regional level. India and Brazil have reported the highest leprosy burden in the last few decades, revealing the need for strategies and tools to help health professionals correctly manage and control the disease.

**Objective:** The main objective of this study was to develop a cross-platform app for leprosy screening based on artificial intelligence (AI) with the goal of increasing accessibility of an accurate method of classifying leprosy treatment for health professionals, especially for communities further away from major diagnostic centers. Toward this end, we analyzed the quality of leprosy data in Brazil on the National Notifiable Diseases Information System (SINAN).

**Methods:** Leprosy data were extracted from the SINAN database, carefully cleaned, and used to build AI decision models based on the random forest algorithm to predict operational classification in paucibacillary or multibacillary leprosy. We used Python programming language to extract and clean the data, and R programming language to train and test the AI model via cross-validation. To allow broad access, we deployed the final random forest classification model in a web app via shinyApp using data available from the Brazilian Institute of Geography and Statistics and the Department of Informatics of the Unified Health System.

**Results:** We mapped the dispersion of leprosy incidence in Brazil from 2014 to 2018, and found a particularly high number of cases in central Brazil in 2014 that further increased in 2018 in the state of Mato Grosso. For some municipalities, up to 80% of cases showed some data discrepancy. Of a total of 21,047 discrepancies detected, the most common was “operational classification does not match the clinical form.” After data processing, we identified a total of 77,628 cases with missing data. The sensitivity and specificity of the AI model applied for the operational classification of leprosy was 93.97% and 87.09%, respectively.

**Conclusions:** The proposed app was able to recognize patterns in leprosy cases registered in the SINAN database and to classify new patients with paucibacillary or multibacillary leprosy, thereby reducing the probability of incorrect assignment by health centers. The collection and notification of data on leprosy in Brazil seem to lack specific validation to increase the quality of the data for implementations via AI. The AI models implemented in this work had satisfactory accuracy across Brazilian states and could be a complementary diagnosis tool, especially in remote areas with few specialist physicians.

(*JMIR Mhealth Uhealth* 2021;9(4):e23718) doi:[10.2196/23718](https://doi.org/10.2196/23718)

**KEYWORDS**

leprosy; artificial intelligence; random forest; Python; R; apps; mHealth; shinyApp

## Introduction

### Leprosy Background

Leprosy is an infectious disease caused by *Mycobacterium leprae* and *Mycobacterium lepromatosis*, which affects the skin and peripheral nerves, causing stigma and disabilities that limit community involvement and social engagement [1]. According to the World Health Organization (WHO) report in 2020, there were 208,619 new leprosy cases registered globally in 2018, with a prevalence rate corresponding to 0.2 in 10,000 individuals. Brazil reported 27,893 new cases in 2019 and is considered the country with the second highest number of new leprosy cases registered in the world. Moreover, Brazil reported the highest number of retreated cases (6887), followed by India (5332). In-depth analysis of retreatment to determine the reasons for treatment interruption would improve compliance, effective use of drugs, and, to some extent, prevent the emergence of drug resistance. A review of new cases detected in the three most highly endemic countries (India, Brazil, and Indonesia) indicated that the number of new cases, the proportion of child cases, and those associated with disability have barely changed in Brazil and Indonesia for the past 5 years. Therefore, there is a long way to go to achieve the following criteria for leprosy elimination: (1) confirmed absence of children with leprosy for 5 consecutive years and (2) confirmed absence of new leprosy cases for 10 years [2].

Currently, the conventional diagnosis of leprosy is typically based on clinical evaluation alone, especially when histopathological analysis is not available. The clinical diagnosis is based on cardinal signs such as the presence of skin lesions (often with loss of sensitivity), thickening of the nerves, and presence of the pathogen in a skin smear or histological tissue samples. Based on this information, classifications are applied to aid in the understanding and treatment of the disease [1,3].

The Madrid classification divides leprosy patients into the following four categories: indeterminate, tuberculoid, borderline, and lepromatous [3]. However, the WHO has also proposed an operational classification to facilitate fieldwork. Patients with up to five skin lesions are classified as having paucibacillary leprosy and those with more than five skin lesions are classified as having multibacillary leprosy [4].

Currently, multidrug therapy is the main treatment for leprosy, which is based on schemes supported by the operational classification. The Guidelines Development Group, established by the WHO in 2018, recommends the same regimen of three drugs (rifampicin, dapsone, and clofazimine) for all leprosy patients, with a 6-month duration for paucibacillary cases and a 12-month duration for multibacillary cases. Some evidence suggests a potential increase in the risk of relapse for patients with paucibacillary leprosy using the previous two-drug regimen [4]. Therefore, this three-drug regimen has the potential to reduce the consequences of misclassifying multibacillary cases as paucibacillary cases (based on lesion count) and the implementation advantages of using the same combination of three drugs for both types [4-6].

Laboratory diagnosis can help differentiate leprosy from other dermatological/neurological diseases, especially in cases of suspected recurrence, and determine an appropriate treatment duration. In these cases, microscopic examination of the dermal smear is the method most commonly used because it is easy to perform and is of low cost. The bacilloscopy index (BI) is negative (0) in the tuberculoid and indeterminate forms, is strongly positive in the lepromatous type, and reveals a variable result in borderline cases [3].

In 2020, the WHO outlined the goal to interrupt leprosy transmission at the national or regional level by 2030. However, to achieve this goal, it is necessary to routinely implement active case detection and contact tracing. Therefore, it is urgent to improve the tools used for an early and precise diagnosis of new cases [2].

### App Characteristics

Brazil uses the Sistema de Informação Nacional de Agravos de Notificação/National Notifiable Diseases Information System (SINAN) to deal with epidemiological aspects of diseases with compulsory notification. A leprosy-specific form has to be filled out for each confirmed case, which involves information about the patient, including the number of lesions and affected nerves, grade of physical disability, and demographic variables, among others. All of these data are stored in SINAN's online database and are available for epidemiological studies [7-10].

The app proposed in this study was based on an in-depth analysis of the SINAN database using machine learning, a research area that focuses on how computers acquire knowledge from data. Machine learning can be subclassified into two general types: unsupervised learning and supervised learning. Unsupervised learning does not have a focus on a predictable output, as its main objective is to identify data patterns. By contrast, supervised learning focuses on an outcome such as determining if an animal in a picture is a cat or a dog [11]. For app development in this study, we used random forest (RF), as one of the most effective algorithms of supervised learning that was developed by Leo Breiman [12], to predict the operational classification (paucibacillary or multibacillary) of a given leprosy case. Our research team previously successfully used this algorithm to predict new cases of leprosy among household contacts using molecular and serological inputs [13].

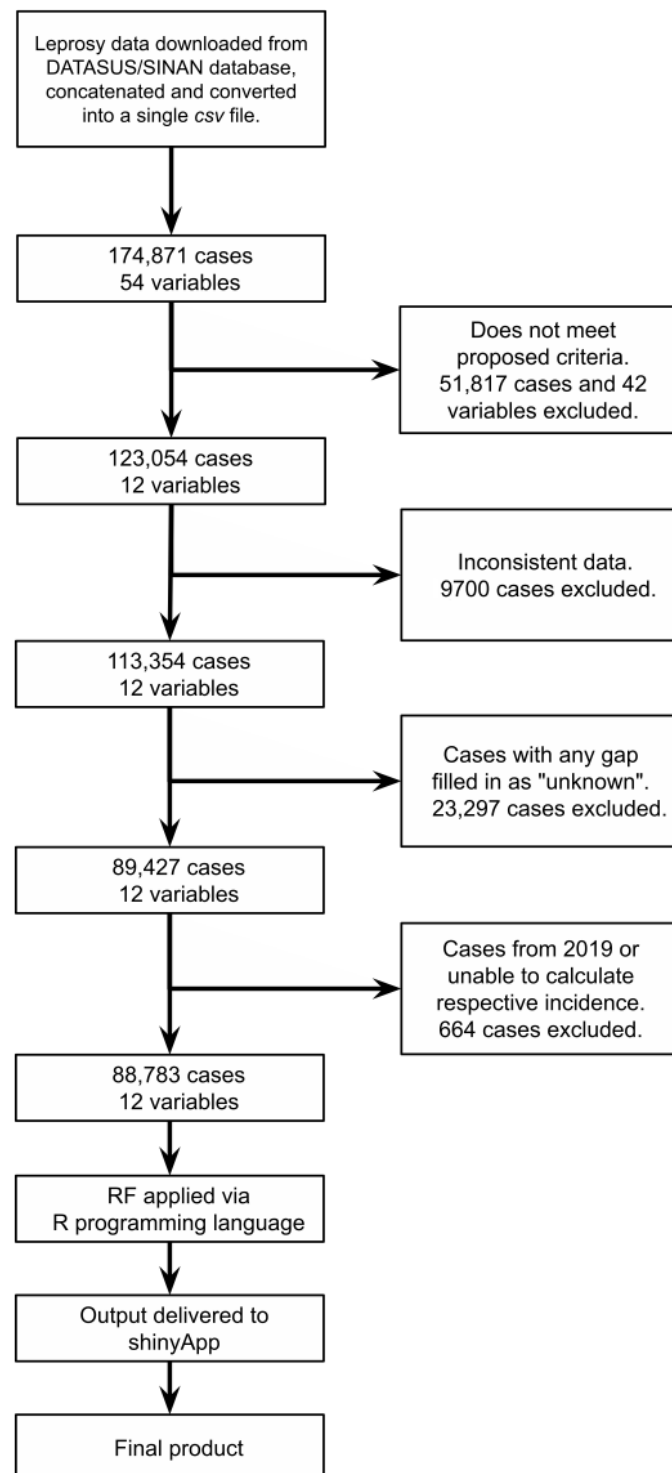
There are some existing apps that were also designed to help diagnose neglected tropical diseases (NTDs). According to the WHO in 2018, an app was developed to facilitate the diagnosis of NTDs of the skin, including leprosy. This app allows health care workers and the public to obtain information about a specific disease, such as its clinical features, management, and geographical distribution, and provides a list of potential diagnoses. The training guide is now updated with recent information and has been translated into an easy-to-use interactive mobile app available free of charge on both Android and iOS platforms [14]. Recently, a new app that helps health care professionals diagnose leprosy was launched by the Federal University of São Paulo [15]. However, among the cited references, none describes the creation process of apps designed to control leprosy using artificial intelligence (AI) to predict

the diagnosis, leaving an important gap in the field that this study aimed to address.

### Study Objective

Given the above background, the main objective of our study was to develop a cross-platform app for leprosy screening based on AI. This app was designed to recognize patterns in leprosy cases registered in the SINAN database and to classify new cases as paucibacillary or multibacillary, thereby reducing the probability of misclassification by the health center.

**Figure 1.** Flow diagram summarizing the data-processing and app-building steps. SINAN: Sistema de Informação Nacional de Agravos de Notificação (National Notifiable Diseases Information System); DATASUS: DATASUS: Sistema Único de Saúde (Unified Health System) data portal; RF: random forest; csv: Comma Separated Value.



## Methods

### Overview

We divided the stages of app construction into two steps: (1) processing raw data and obtaining a decision matrix, and (2) using the decision matrix to build the app for classifying a case given an input. An overview of the entire process is shown in Figure 1.

### Processing Raw Data

Initially, we downloaded all SINAN records related to leprosy cases from 2014 to 2019, which were converted to a single Comma Separated Values file. This procedure resulted in a 54-column file containing data on 174,871 cases, with each

column corresponding to a specific variable reported by Brazilian health professionals about the leprosy cases notified.

Many of these columns (variables) were not relevant to our study. Therefore, we removed those that did not fulfill the criteria as shown in [Table 1](#).

**Table 1.** Exclusion criteria and justifications.

Exclusion criteria	Justifications
Columns with more than 25,000 “NA” (not available)	The objective was to remove variables that many professionals have not declared the value of, as a large amount of missing data may impair processing. The number 25,000 was arbitrarily defined, focusing on not drastically reducing the total amount of data
Categorical variable with more than 53 input possibilities	R shows an alert when a categorical variable with more than 53 input possibilities is being used, given that the greater the number of input possibilities, the smaller the meaning of each input to the model
Variables that may induce a result	Some variables imply an operational classification (eg, “g-MB” therapeutic scheme implies that the patient has a case of multibacillary leprosy), causing bias to the model
Variables with no apparent correlation with the prediction.	Boruta [16], an algorithm to find the most relevant variables to predict outcomes in a given dataset, was used; variables with little relevance were excluded
Redundant variables	Redundant variables do not provide additional information to the model, and therefore there is no reason to keep both. An analysis using Python showed that some variables had almost 100% correspondence with another (eg, the state where the case was notified and the state where the patient lives). The Boruta algorithm is also useful to remove redundant variables.

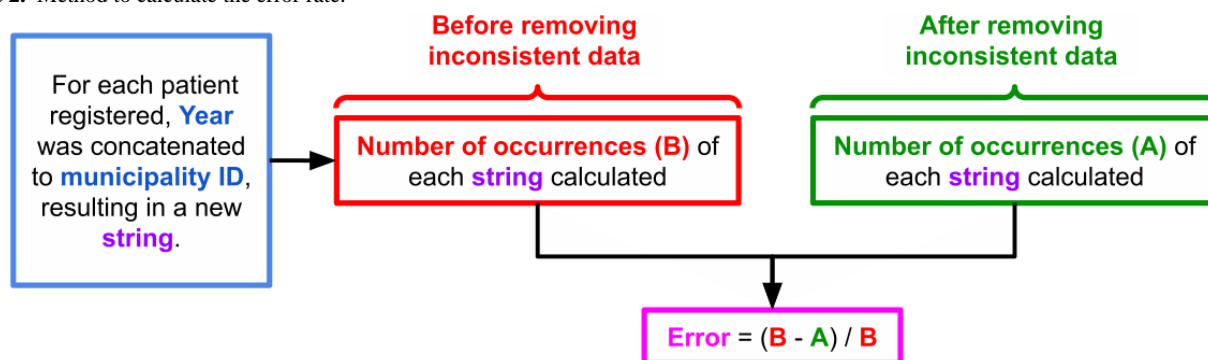
The remaining dataset was composed of the variables age, gender, race, education, grade of disability, operational classification, BI, number of affected nerves, clinical form, municipality ID, number of household contacts, and the number of skin lesions. After this processing, we removed all lines with any entry of “NA” (not applicable), leaving a total of 123,054 cases.

with these criteria, 113,354 cases remained. All lines with missing items (gender, race, schooling, grade of disability, operational classification, bacilloscopy, or clinical form) were removed to improve the accuracy of prediction, leaving 89,427 cases.

The Brazilian Practical Guide on Leprosy was reviewed to define the following criteria to remove cases with any inconsistency: (i) samples with a positive BI are always cases of multibacillary leprosy; (ii) patients with paucibacillary leprosy should have five or fewer skin lesions; (iii) indeterminate and tuberculoid are always paucibacillary forms; (iv) borderline and lepromatous cases are always multibacillary forms; (v) indeterminate cases have no disability; and (vi) a maximum of approximately 18 nerve trunks are evaluated in clinical examinations [17]. After removing cases that were not consistent

The Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics), responsible for conducting the census, provides tables containing the estimated population by year and by municipality [18]. These tables were merged with all SINAN data (174,871 cases) to calculate the leprosy incidence by year and by city. However, it was necessary to delete the data for 2019 because of the lack of year-round records in SINAN, leaving 88,783 cases. Finally, comparing the number of cases after removing inconsistencies, we created an index that represents the percentage of patients with discrepancies in diagnosis (error) per year for each municipality, according to the workflow shown in [Figure 2](#).

**Figure 2.** Method to calculate the error rate.



Confidence intervals of these municipality inconsistencies were calculated for each state of Brazil. We calculated the median of household contacts by city and year. All data processing up to this point was performed using Python 3.8 and WPS spreadsheets.

### Applying RF and Building the App

After initial processing with Python, the RF algorithm was applied to the resulting data using the R software package Random Forest. In addition to RF, there are several other machine-learning classification algorithms that could be appropriate for this task, such as naive Bayes [19], logistic regression [20], k-nearest neighbor [21], decision tree [22], and gradient boosting [23]. This list is not exhaustive, but includes the most common algorithms that were applied to our dataset for model comparison. The RF algorithm was ultimately chosen owing to its better performance, ease of use and interpretation, and wide dissemination in the literature. RF is an ensemble learning method for classification, regression, and other tasks, which operates by constructing a multitude of decision trees during training and then outputting the class that represents the mode of the categories (classification) or mean prediction (regression) of the individual trees [24].

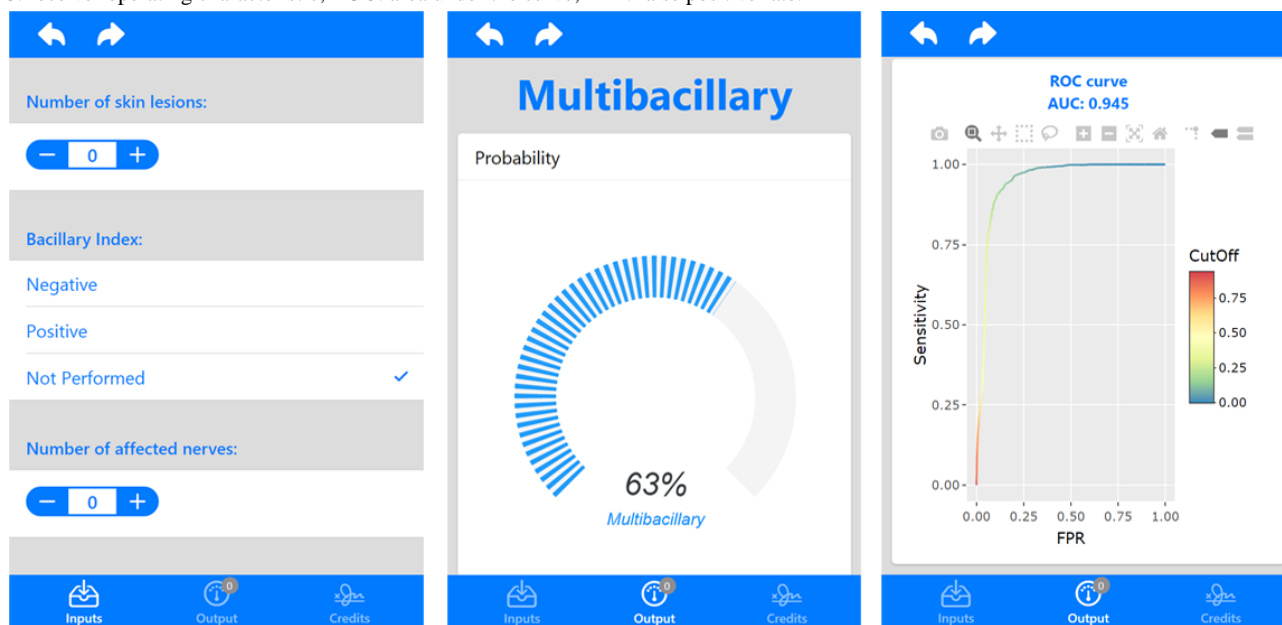
After several tests to improve model accuracy, the following subset of variables was used to predict the operational classification of each case: region, state, city, age, number of skin lesions, affected nerves, household contacts, and

bacilloscopy. Prediction of operational classification was chosen as the metric for evaluation instead of prediction of the clinical form for two main reasons: (1) Brazilian treatment is based on the operational classification [17], and (2) dichotomous models (in this case paucibacillary or multibacillary) have better results in RF-based analyses [25].

This multiplatform app was designed to meet the scientific demand for technological innovation concurrently with the lack of safe and accurate diagnoses in remote regions of Brazil, where training in clinical practice does not always match the international standards recommended by the WHO. In this sense, we incorporated only clinical variables reported by SINAN so that the app would be useful in the Sistema Único de Saúde (SUS; Unified Health System) throughout Brazil.

The decision forest obtained by the RF algorithm from the R package was deployed in a ShinyApp environment, which is a web service for constructing a friendly user interface. Figure 3 shows the structure and flow of the app. The receiver operating characteristic (ROC) curve [26] represents the true positive rate (TPR or sensitivity) as a function of the false positive rate (FPR or 1 – specificity). The colored scale on the right of the curve denotes the cutoffs distribution used to obtain a classification of multibacillary leprosy by the model. Each cutoff is related to one TPR and FPR point in the plot [27]. Given the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), the TPR and FPR are defined as  $TP/TP+FN$  and  $FP/FP+TN$ , respectively.

**Figure 3.** Screenshot representing the R ShinyApp input and output flows. The layout of the app may eventually change to improve user experience. ROC: receiver operating characteristic; AUC: area under the curve; FPR: false positive rate.



A good model must enhance the TPR and decrease the FPR. Thus, the quality of the model may be represented by the area under the ROC curve (AUC) value [28]. The best models have an AUC close to 1 and the worst models have an AUC close to 0.

The value of each variable in the database has a different weight for the model. Some values approximate a paucibacillary classification, whereas other values approximate toward a

multibacillary classification. Representing the distance between paucibacillary and multibacillary as a scale from 0 to 1, we can choose different values on this scale as the limit between the two classifications. These possible values are represented by the colored scale to the right of the ROC curve in Figure 3.

## Results

### Database Processing

#### Inconsistencies

Table 2 shows the 21,053 inconsistencies obtained in the dataset from a total of 174,871 individuals in the SINAN database in

**Table 2.** Number of occurrences per inconsistency.

Inconsistency	Number of occurrences
Operational classification does not match the clinical form	8545
Indeterminate with disability	4867
Indeterminate with affected nerves	3785
Paucibacillary with positive bacilloscopy	2825
Paucibacillary with more than 5 skin lesions	938
Patients with more than 18 affected nerves	93

#### Missing Data

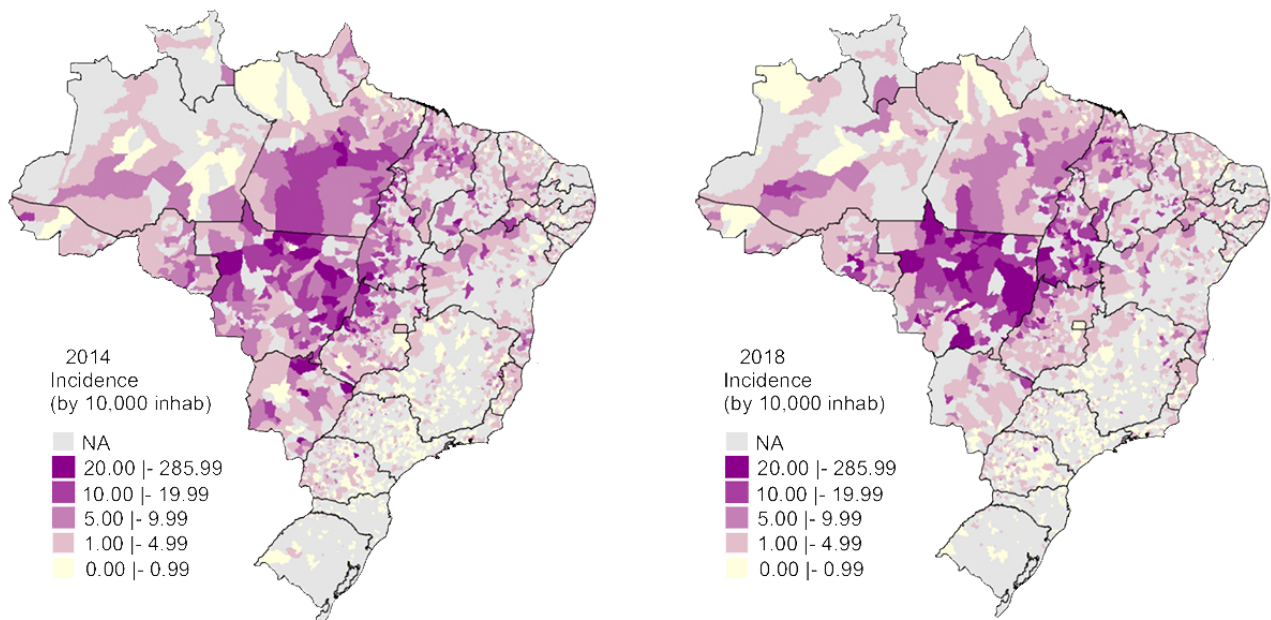
Before data processing, the SINAN database had 35,616 lines with at least one NA, 26,539 lines with at least one unknown item (gender, race, schooling, grade of disability, operational classification, bacilloscopy, or clinical form), and 15,473 lines with both NA and an unknown item. There were a total of 77,628 cases with missing data, accounting for 44.39% of the total 174,871 cases.

the period of 2014-2018. It is important to note that an individual might have more than one discrepancy. The main inconsistency identified was “operational classification does not match the clinical form.” This means that a lepromatous or borderline patient was misclassified as a paucibacillary patient, or that an indeterminate or tuberculoid patient was misclassified as a multibacillary patient.

### Leprosy in Brazil

Based on available data in SUS, after cleaning the dataset, it was possible to geographically visualize the dispersion of leprosy incidence in Brazil over the period from 2014 to 2018 (Figure 4). A high number of cases were notably located in central Brazil. This number became even higher in 2018, highlighting the State of Mato Grosso with the highest number of leprosy case reports in the country. Some municipalities had an incidence of up to 285 cases per 10,000 inhabitants.

**Figure 4.** Geographic distribution of new annual cases of leprosy in Brazilian municipalities. inhab: inhabitants; NA: not available.



In addition to showing the geographical extent of the disease, the cleaned database included individuals from 4 to 106 years old (mean of 44 years), with an average of seven lesions, two affected nerves, and three household contacts per case. The BI was not calculated in 37.00% (32,853/88,783) of cases, and for

the remaining cases, 41.18% (23,034/55,930) of the BI results were positive. In total, the database reported multibacillary leprosy in 76.66% (68,061/88,783) of patients, who were scattered throughout the Brazilian territory (Figure 5).

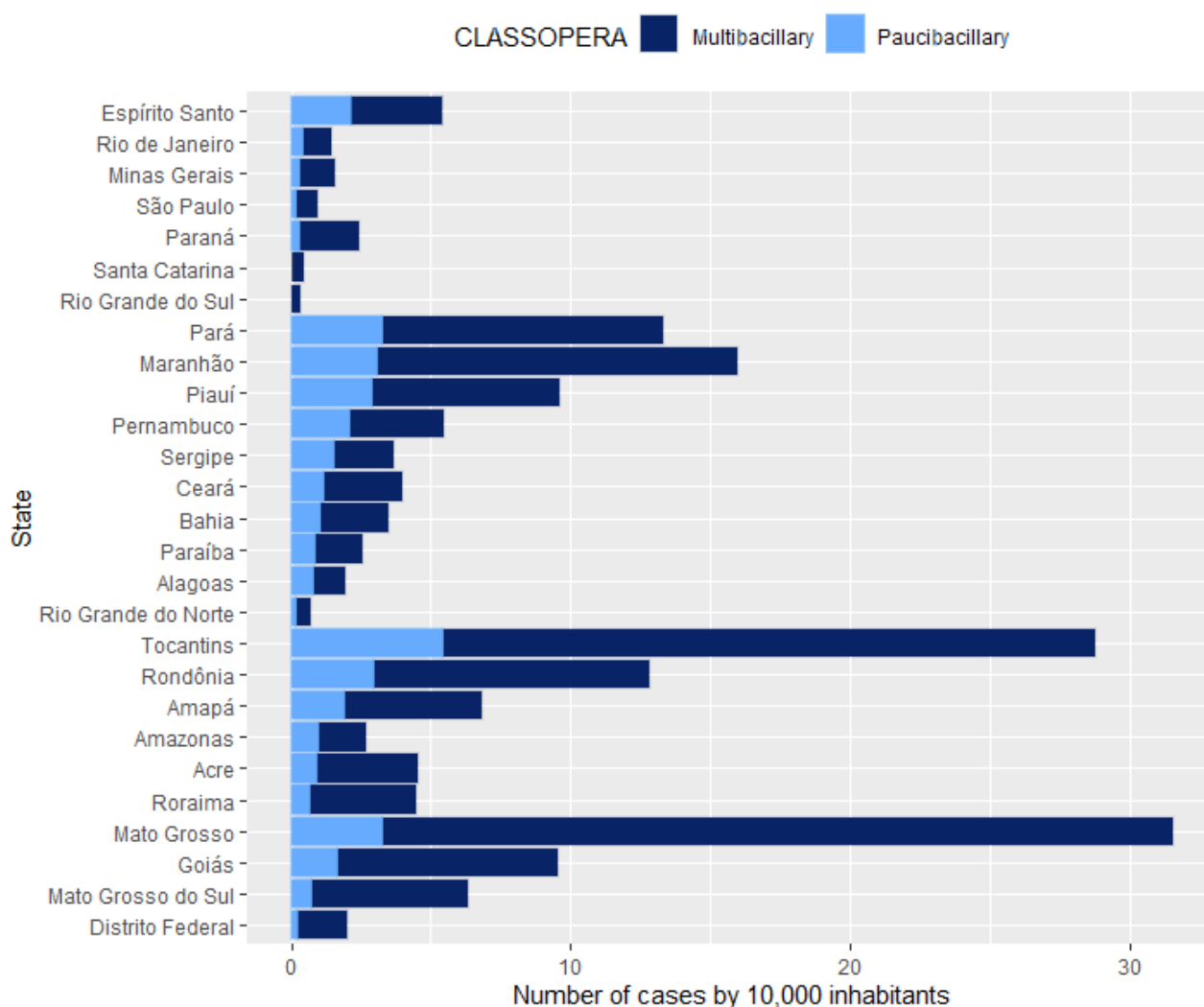
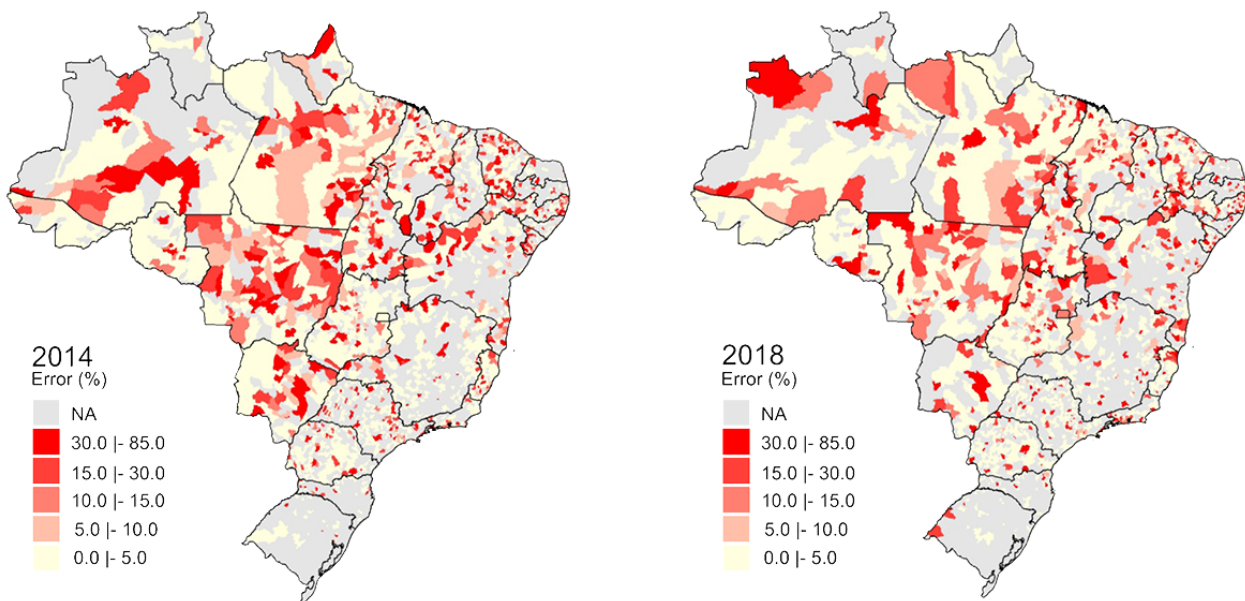
**Figure 5.** Distribution of leprosy cases in the Brazilian states from 2014 to 2018.

Figure 5 shows the distribution of paucibacillary and multibacillary cases among Brazilian states (per 10,000 inhabitants) with the highest numbers found in Mato Grosso (31.5), Tocantins (28.8), Maranhão (15.9), Pará (13.3), and Rondônia (12.8), which are all states with a predominance of multibacillary cases.

Figure 6 shows that leprosy misclassification exhibited a certain degree of homogeneity throughout Brazil from 2014 to 2018,

which does not raise a suspicion of its correlation with leprosy's incidence. Notably, some municipalities had discrepancy rates reaching up to 80%, although in general this error rate appears to have decreased over time. Unlike other Brazilian states, Amazonas (in 2018) presented an unreasonable increase in the error rate, especially considering that these municipalities had virtually no notification of the disease in 2014.

**Figure 6.** Geographic leprosy misclassification distribution in Brazilian municipalities. NA: not available.

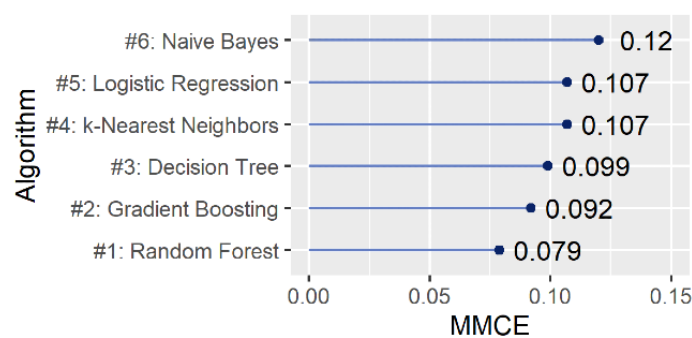


**AI Model to Support Clinical Diagnosis**

Starting from the assumption that the error of clinical diagnosis is properly characterized, it was possible to develop a support model based on AI. The RF algorithm presented the smallest

mean of misclassification error compared with the other algorithms using 10-fold cross-validation with default hyperparameters considered in the mlr R package [29] (Figure 7).

**Figure 7.** Comparison of algorithms according to the mean of misclassification error (MMCE).



We next sought to determine if AI can assist in choosing the correct treatment for leprosy. Since the Brazilian Ministry of Health and the WHO suggest patterns (an algorithm) to classify the disease, our group was able to develop this new strategy to help control leprosy. A different model for each Brazilian state was used to improve the prediction in the remaining states; that is, a cross-validation strategy was applied to avoid overfitting by training in one state but testing in others. The number of lesions, incidence, and affected nerves were among the most

important variables in the three best models (Mato Grosso, Rio Grande do Sul, and Paraná).

Figure 8 shows a heatmap of the quality of all models for each state of Brazil. The blue scale represents the accuracy of the models in each testing situation. The Y axis denotes the states used to build the models (training dataset) and the X axis indicates the states where the models were applied (testing dataset).



**Figure 8.** Quality in the classification of leprosy cases by artificial intelligence models in Brazilian states.

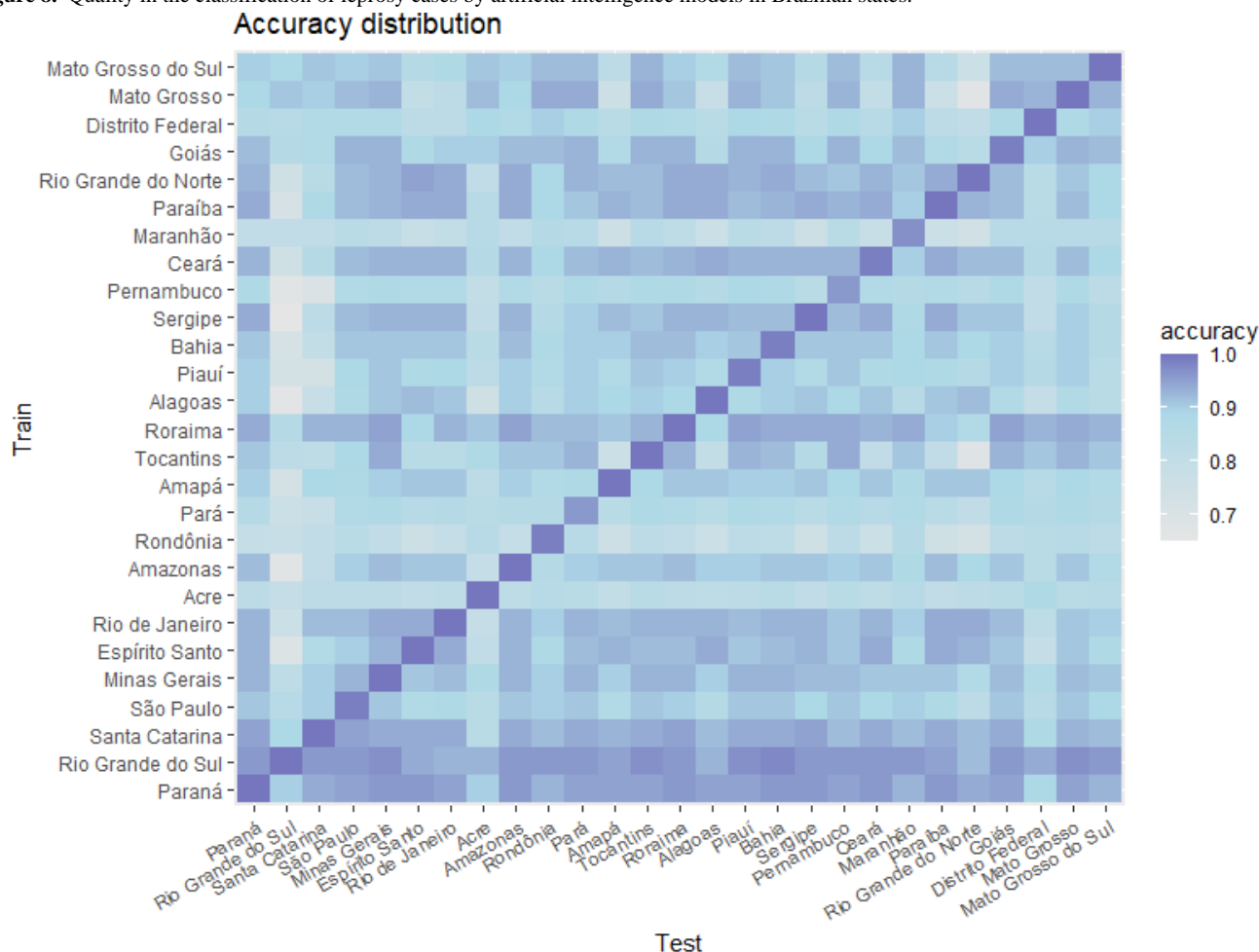


Table 3 shows the importance of the three best models (Mato Grosso, Rio Grande do Sul, and Paraná) used in development of the app. These models showed some variability in the importance of each variable. In the Mato Grosso model, the most important variables were the number of affected nerves, number of skin lesions, and incidence. In the Rio Grande do

Sul model, the main variables were number of skin lesions and number of affected nerves. Finally, in the Paraná model, the most important variables were the number of skin lesions and bacilloscopy. “Gender” represented less than 2% of importance in all three models, whereas “number of skin lesions” represented more than 23% importance in all models.

**Table 3.** Importance (in percent) of each variable utilized in the models that represent the highest accuracy.

Variable	Meaning	Mato Grosso model	Rio Grande do Sul model	Paraná model
INCIDÊNCIA	Incidence	15.0	9.5	6.1
NU_IDADE_N	Age	9.3	8.1	5.5
CS_SEXO	Gender	1.6	1.7	1.9
CS_RACA	Race	2.7	6.9	1.2
CS_ESCOL_N	Educational level	4.7	6.4	3.4
NU_LESOES	Number of skin lesions	23.6	37.0	41.9
AVALIA_N	Grade of disability	4.4	4.2	5.0
BACILOSCOP	Bacilloscopy	6.1	4.5	23.7
CONTREG	Number of household contacts	5.1	5.9	2.8
NERVOSAFET	Number of affected nerves	27.5	15.8	8.5

Table 4 shows the quality of the AI model applied to the differential diagnoses of paucibacillary and multibacillary leprosy in 26,546 cases. We considered multibacillary as the

reference. Some measures such as accuracy (proportion of multibacillary and paucibacillary cases correctly classified), sensitivity (proportion of correctly classified cases given that

they were truly multibacillary), specificity (proportion of correctly classified cases given that were truly paucibacillary), positive predictive value (proportion of multibacillary cases given a positive classification by the model), and negative

predictive value (proportion of paucibacillary cases given a negative classification by the model) are presented for each model.

**Table 4.** Quality of the artificial intelligence model applied to the differential diagnosis of paucibacillary and multibacillary leprosy.

Quality parameter	Mato Grosso model	Rio Grande do Sul model	Paraná model
Accuracy	0.970	0.812	0.929
Sensitivity	0.926	0.977	0.877
Specificity	0.812	0.218	0.919
PPV <sup>a</sup>	0.936	0.803	0.972
NPV <sup>b</sup>	0.786	0.740	0.698

<sup>a</sup>PPV: positive predictive value.

<sup>b</sup>NPV: negative predictive value.

## Discussion

### Relevance of Correct Operational Classification

This study analyzed the SINAN database from 2014 to 2018 considering the epidemiological, clinical, and sociodemographic context of patients diagnosed with leprosy in Brazil. In analyzing the frequency of inconsistencies in the SINAN database (Table 2), we observed that the main disagreement was operational classification (ie, multibacillary vs paucibacillary) that did not match the clinical form. These results highlight the importance of training health professionals to correctly identify the operational class and improving data collection. Information on the operational classification for a given case is essential to select the appropriate treatment scheme. Misclassification may harm both the patient and the health care system. A patient with paucibacillary leprosy receives treatment for 6 months, whereas a patient with multibacillary leprosy receives treatment for 12 months [17,30]. Therefore, a patient with paucibacillary leprosy misclassified as multibacillary leprosy would undergo 6 months of unnecessary treatment, burdening the health care system and increasing the risk of suffering from side effects during this time. By contrast, a patient with multibacillary leprosy misclassified as paucibacillary leprosy could be undertreated, requiring restarting the treatment that leads to spending more resources with potential subsequent disability rehabilitation.

According to Grossi et al [31], there is a strong tendency for health professionals to classify patients as having multibacillary leprosy, at least in Brazil. This trend, confirmed in our analysis of the SINAN database, seemed to be related to the absence of laboratory tests such as skin biopsy and slit-skin smear that provide security to professionals in decision making. Furthermore, Nobre et al [32] showed that the proportion of newly diagnosed leprosy cases that were cases of multibacillary leprosy increased by 11.6% during a study of cases reported in Brazil from 2001 to 2013. It is important to keep in mind that leprosy classification methodology may vary according to the diagnostic capacity of the center. In primary care settings, diagnosis and operational classification are mainly based on clinical findings, whereas in reference centers, confirmatory biopsies can be performed. Moreover, it is important to note

that improper treatment can also lead to bacterial resistance [30-33].

### Relevance of the App

According to WHO goals for 2030, it is necessary to employ strategic methodologies to assist leprosy control. The use of AI is a novel method with potential to expand the capacities to diagnose diseases, especially those that are neglected. The use of AI therefore allows for obtaining higher coverage in the initial diagnosis process and facilitates the sharing of secure information, with the aim to expand and reach a larger number of health professionals [7].

We recognize the importance of a tool to improve the accuracy of insufficiently trained health professionals, especially in the most remote areas of Brazil. The app presented in this work proved to be a promising option to improve the coverage and scalability to the Brazilian health service regarding the choice of an appropriate treatment for leprosy [33]. The training of a physician requires considerable time when compared to computational diagnostic resources. In addition, the accessibility of this app in the hands of any professional via mobile or desktop devices offers scalability that traditional teaching methods cannot achieve [34]. Machine learning in standardized functions or decisions is expected to be much faster than the acquisition of general human knowledge. Therefore, our app would provide speed, scalability, and broadcasting to fight leprosy without compromising accuracy [35].

### Appropriate Data Collection for Leprosy Classification

Another relevant issue to mention involves problems in incorrectly filling out the reporting form. To fill out a SINAN form correctly, it is necessary to list the Madrid Classification and the treatment according to the operational classification. Therefore, a patient given a classification of indeterminate or tuberculoid has to receive treatment for paucibacillary leprosy, whereas a classification of borderline or lepromatous requires treatment for multibacillary leprosy.

According to the Brazilian Ministry of Health in 2017, a case with a positive BI result should be considered as multibacillary leprosy. However, doubts arise for cases that are considered to be borderline and close to the tuberculoid pole. Despite the

difficulty of correct classification, these cases are generally considered to be multibacillary leprosy [17]. Another important point is related to the high rate of patients with missing or unknown data in the SINAN database. There were 77,628 patients with at least one data item missing, which represents 44% of the total. These data could be essential to improve the accuracy of machine-learning models in which performance is related to the amount of available clean data. This issue reinforces the necessity for developing a better approach for inputting data to the SINAN database as well as adopting good practices for validating these data. Lastly, the part of the form related to clinical aspects is filled out by the doctor, whereas the first part related to patient identification, in a general way, is filled out by the auxiliary staff. The state of Minas Gerais adopted this criterion; however, in other areas, other health professionals might complete the entire SINAN form, even the parts that are supposed to be completed by the physician. The Ministry of Health recommends that the physicians fill out the SINAN, make the notification and the diagnosis, as well as the documentation of successful treatment [17].

### Relevance of Using AI

In addition to the implementation of an algorithm that helps choose the correct therapy, the use of intelligent data collection devices would allow for higher quality and validation [36]. Such devices could be used to guarantee higher-quality data for future research and innovations regarding the improvement of health services as a whole, since poor-quality data is not limited to leprosy [37]. An algorithm is nothing more than a finite sequence of well-defined instructions that can be automated, usually to solve problems. The actions advocated by the WHO for the classification of leprosy to ensure appropriate treatment and cure of this disease represent an example of the algorithm itself. Considering this evident algorithm, our mission is to provide a statistically reliable emulation to aid in the diagnosis of leprosy. We have implemented a classification method based

on RF, which is accessible using any and all devices with internet access.

Notably, the high accuracy (92.38%), sensitivity (93.97%), and specificity (87.09%) of this app provide a multiplatform method to support scalable characterization/classification for numerous other neglected diseases in remote communities in Brazil and worldwide.

### Limitations

As previously mentioned, SINAN is a platform launched to manage notifications from each Brazilian state. The records of this platform are obtained from assessments made by many health professionals with different levels of qualification. Thus, the quality of data depends on many factors, including (i) quality of the items requested by the forms and their correct interpretation, (ii) correct clinical assessment of the patient, and (iii) proper filling out of the forms. In addition, it is important to note that the possibility to add more items of information about serological and molecular integrated tests for leprosy diagnoses could undoubtedly improve the accuracy of the method, as we have done in our research group [13]. Lastly, the app is not currently available without an internet connection since the AI model is deposited in the RStudio cloud.

### Conclusions

The proposed app showed good accuracy to classify a case as paucibacillary or multibacillary leprosy by recognizing patterns in leprosy cases registered in the SINAN database. After validation, this app could be an essential tool to help health professionals make an accurate leprosy operational classification and decide which treatment to use for patients with paucibacillary or multibacillary leprosy considering reducing the likelihood of mistreatment. This study also highlights the importance of improving data collection methods given that prediction accuracy markedly increases with improved data quality.

---

### Acknowledgments

The authors thank Artur José Vilar Sette, Davi Metzker Júnior, and Vladimir Machado Rios for helping in some aspects of app building, and Tillman Rauh for helping to translate some parts of this article. We are also grateful to all members of CREDEN-PES, Programa Multicêntrico de Bioquímica e Biologia Molecular at Universidade Federal de Juiz de Fora Campus Governador Valadares, and to PROEX/PROPP/UFJF. This study received financial support from the Conselho de Desenvolvimento Tecnológico e Científico/CNPq/BRAZIL, FAPEMIG. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES; Finance Code 001, file number: 88881.361990/2019-01 [Migrated-SICAPES3]). The funding sources had no role in the design of the study; in the collection, analysis, implementation, and interpretation of data; and in writing the manuscript.

---

### Authors' Contributions

GAL processed all raw data from SINAN using Python 3.8 and WPS spreadsheets. MLMDS used these data to apply the RF algorithm in R software and programmed the app (via ShinyApp). GAL and MLMDS prepared the figures and, along with LADOF have authorized, reviewed, and edited this article. ACB and JKF provided expertise for interpretation of the results and critically edited the manuscript. All authors have read and approved the final version of the manuscript.

---

### Conflicts of Interest

None declared.

---

## References

1. Ahuja M, Lavana M, Singh I, Turankar RP, Chhabra S, Narang T, et al. Detection of Mycobacterium lepromatosis in patients with leprosy in India. *Infect Drug Resist* 2018 Oct;11:1677-1683. [doi: [10.2147/IDR.S166035](https://doi.org/10.2147/IDR.S166035)] [Medline: [30349325](https://pubmed.ncbi.nlm.nih.gov/30349325/)]
2. Global leprosy (Hansen disease) update, 2019: time to step-up prevention initiatives. *World Health Organization Weekly Epidemiological Record*. 2019. URL: <https://www.who.int/publications/i/item/who-wer9536> [accessed 2021-03-26]
3. Araújo MG. Hanseníase no Brasil. *Rev Soc Bras Med Trop* 2003 Jun;36(3):373-382. [doi: [10.1590/s0037-86822003000300010](https://doi.org/10.1590/s0037-86822003000300010)]
4. World Health Organization. Guidelines for the diagnosis, treatment and prevention of leprosy. New Delhi: World Health Organization Regional Office for South-East Asia; 2018.
5. Lastória JC, Abreu MAMM. Hanseníase: diagnóstico e tratamento. *Diagn Tratamento* 2012;17(4):173-179 [FREE Full text]
6. Ministério da Saúde Brasil. Guia de Vigilância em Saúde: volume único. Brazil: Editora do Ministério da Saúde; 2019.
7. World Health Organization. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021-2030: overview. Geneva: World Health Organization; 2020.
8. Ministério da Saúde Brasil. Ministério da Saúde. In: Sistema de Informação de Agravos de Notificação – SINAN: normas e rotinas. Brazil: Editora do Ministério da Saúde; 2007.
9. Arquivos de Dados. Departamento de Informática do SUS (DATASUS). URL: <http://www2.datasus.gov.br/DATASUS/index.php?area=0901> [accessed 2019-07-25]
10. SINAN database. GitHub. URL: [https://github.com/gbayres/SINAN\\_database](https://github.com/gbayres/SINAN_database) [accessed 2021-03-25]
11. Deo RC. Machine learning in medicine. *Circulation* 2015 Nov 17;132(20):1920-1930 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.115.001593](https://doi.org/10.1161/CIRCULATIONAHA.115.001593)] [Medline: [26572668](https://pubmed.ncbi.nlm.nih.gov/26572668/)]
12. Breiman L. Random forests. *Machine Learn* 2001;45:5-32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
13. Gama RS, Souza MLM, Sarno EN, Moraes MO, Gonçalves A, Stefani MMA, et al. A novel integrated molecular and serological analysis method to predict new cases of leprosy amongst household contacts. *PLoS Negl Trop Dis* 2019 Jun;13(6):e0007400 [FREE Full text] [doi: [10.1371/journal.pntd.0007400](https://doi.org/10.1371/journal.pntd.0007400)] [Medline: [31181059](https://pubmed.ncbi.nlm.nih.gov/31181059/)]
14. Neglected tropical diseases of the skin: WHO launches mobile application to facilitate diagnosis. World Health Organization. 2020 Jul 16. URL: <https://www.who.int/news/item/16-07-2020-neglected-tropical-diseases-of-the-skin-who-launches-mobile-application-to-facilitate-diagnosis> [accessed 2021-03-26]
15. Aplicativo auxilia profissionais da saúde no diagnóstico de hanseníase. *Jornal da USP*. 2020 Mar 09. URL: <https://jornal.usp.br/?p=304921> [accessed 2021-03-26]
16. Kurska MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Soft* 2010;36(11):1-13 [FREE Full text] [doi: [10.18637/jss.v036.i11](https://doi.org/10.18637/jss.v036.i11)]
17. Ministério da Saúde Brasil. Guia prático sobre a hanseníase. Brazil: Editora do Ministério da Saúde; 2017.
18. Estimativas da População. Instituto Brasileira de Geografia e Estatística (IBGE). URL: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=downloads> [accessed 2021-03-26]
19. Xu W, Jiang L, Yu L. An attribute value frequency-based instance weighting filter for naive Bayes. *Journal Exp Theor Artif Intell* 2018 Nov 09;31(2):225-236. [doi: [10.1080/0952813X.2018.1544284](https://doi.org/10.1080/0952813X.2018.1544284)]
20. Kleinbaum DG, Klein M. Logistic regression: a self-learning text. New York: Springer; 2010.
21. Keller JM, Gray MR, Givens JA. A fuzzy K-nearest neighbor algorithm. *IEEE Trans Syst Man Cybern* 1985 Jul;SMC-15(4):580-585. [doi: [10.1109/tsmc.1985.6313426](https://doi.org/10.1109/tsmc.1985.6313426)]
22. Friedl M, Brodley C. Decision tree classification of land cover from remotely sensed data. *Remote Sens Environ* 1997 Sep;61(3):399-409. [doi: [10.1016/s0034-4257\(97\)00049-7](https://doi.org/10.1016/s0034-4257(97)00049-7)]
23. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal* 2002 Feb;38(4):367-378. [doi: [10.1016/s0167-9473\(01\)00065-2](https://doi.org/10.1016/s0167-9473(01)00065-2)]
24. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002 Dec. URL: [https://www.r-project.org/doc/Rnews/Rnews\\_2002-3.pdf](https://www.r-project.org/doc/Rnews/Rnews_2002-3.pdf) [accessed 2020-03-25]
25. Hanewinkel M, Zhou W, Schill C. A neural network approach to identify forest stands susceptible to wind damage. *For Ecol Manag* 2004 Jul;196(2-3):227-243. [doi: [10.1016/j.foreco.2004.02.056](https://doi.org/10.1016/j.foreco.2004.02.056)]
26. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatr* 2011 Apr 15;48(4):277-287. [doi: [10.1007/s13312-011-0055-4](https://doi.org/10.1007/s13312-011-0055-4)] [Medline: [21532099](https://pubmed.ncbi.nlm.nih.gov/21532099/)]
27. Pepe MS, Cai T, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* 2006 Mar;62(1):221-229. [doi: [10.1111/j.1541-0420.2005.00420.x](https://doi.org/10.1111/j.1541-0420.2005.00420.x)] [Medline: [16542249](https://pubmed.ncbi.nlm.nih.gov/16542249/)]
28. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *Can J Emerg Med* 2006 Jan 21;8(1):19-20. [doi: [10.1017/s1481803500013336](https://doi.org/10.1017/s1481803500013336)] [Medline: [17175625](https://pubmed.ncbi.nlm.nih.gov/17175625/)]
29. Bischl B, Lang M, Kotthoff L, Schiffner J, Richter J, Studerus E, et al. mlr: Machine Learning in R. *J Machine Learn Res* 2016;17:1-5 [FREE Full text]
30. Ramos-e-Silva M, Rebello PF. Leprosy: recognition and treatment. *Am J Clin Dermatol* 2001;2(4):203-211. [doi: [10.2165/00128071-200102040-00001](https://doi.org/10.2165/00128071-200102040-00001)] [Medline: [11705247](https://pubmed.ncbi.nlm.nih.gov/11705247/)]

31. Grossi M, Leboeuf MAA, Andrade ARC, Lyon S, Antunes C, Bühner-Sékula S. The influence of ML Flow test in leprosy classification. *Rev Soc Bras Med Trop* 2008;41(Suppl 2):34-38 [FREE Full text] [doi: [10.1590/s0037-86822008000700008](https://doi.org/10.1590/s0037-86822008000700008)] [Medline: [19618073](https://pubmed.ncbi.nlm.nih.gov/19618073/)]
32. Nobre ML, Illarramendi X, Dupnik KM, Hacker M, Nery J, Jerônimo SMB, et al. Multibacillary leprosy by population groups in Brazil: Lessons from an observational study. *PLoS Negl Trop Dis* 2017 Feb;11(2):e0005364 [FREE Full text] [doi: [10.1371/journal.pntd.0005364](https://doi.org/10.1371/journal.pntd.0005364)] [Medline: [28192426](https://pubmed.ncbi.nlm.nih.gov/28192426/)]
33. Aldo VW, Junior CLB, Wagner HM, Cavalcante C. Ways to implement large scale telemedicine: The Santa Catarina Experience. *Latin Am J Telehealth* 2009;1:364-377.
34. Atienza AA, Patrick K. Mobile health: the killer app for cyberinfrastructure and consumer health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S151-S153. [doi: [10.1016/j.amepre.2011.01.008](https://doi.org/10.1016/j.amepre.2011.01.008)] [Medline: [21521588](https://pubmed.ncbi.nlm.nih.gov/21521588/)]
35. Boulos M, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. *BioMed Eng OnLine* 2011;10(1):24 [FREE Full text] [doi: [10.1186/1475-925x-10-24](https://doi.org/10.1186/1475-925x-10-24)]
36. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *NPJ Digit Med* 2019;2:38-39. [doi: [10.1038/s41746-019-0111-3](https://doi.org/10.1038/s41746-019-0111-3)] [Medline: [31304384](https://pubmed.ncbi.nlm.nih.gov/31304384/)]
37. Santos D, Faria LV, Nascimento AFSC. Philosophical considerations in health: conceptualizing to educate—a perspective on neglected tropical diseases in Brazil. *Creat Educ* 2019;10(06):1125-1139. [doi: [10.4236/ce.2019.106085](https://doi.org/10.4236/ce.2019.106085)]

## Abbreviations

**AI:** artificial intelligence

**AUC:** area under the receiver operating characteristic curve

**BI:** bacilloscopy index

**FN:** false negative

**FP:** false positive

**FPR:** false positive rate

**NA:** not applicable/not available

**NTD:** neglected tropical disease

**RF:** random forest

**ROC:** receiver operating characteristic

**SINAN:** Sistema de Informação Nacional de Agravos de Notificação (National Notifiable Diseases Information System)

**SUS:** Sistema Único de Saúde (Unified Health System)

**TN:** true negative

**TP:** true positive

**TPR:** true positive rate

**WHO:** World Health Organization

*Edited by L Buis; submitted 21.08.20; peer-reviewed by J Ropero, KL Mauco; comments to author 06.10.20; revised version received 01.12.20; accepted 05.01.21; published 07.04.21.*

*Please cite as:*

*De Souza MLM, Lopes GA, Branco AC, Fairley JK, Fraga LADO*

*Leprosy Screening Based on Artificial Intelligence: Development of a Cross-Platform App*

*JMIR Mhealth Uhealth* 2021;9(4):e23718

URL: <https://mhealth.jmir.org/2021/4/e23718>

doi: [10.2196/23718](https://doi.org/10.2196/23718)

PMID: [33825685](https://pubmed.ncbi.nlm.nih.gov/33825685/)

©Márcio Luís Moreira De Souza, Gabriel Ayres Lopes, Alexandre Castelo Branco, Jessica K Fairley, Lucia Alves De Oliveira Fraga. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 07.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Shaping Workflows in Digital and Remote Diabetes Care During the COVID-19 Pandemic via Service Design: Prospective, Longitudinal, Open-label Feasibility Trial

Katarina Braune<sup>1,2</sup>, MD; Karina Boss<sup>1</sup>, CDE, MA; Jessica Schmidt-Herzel<sup>1</sup>, CDE; Katarzyna Anna Gajewska<sup>3</sup>, PhD; Axel Thieffry<sup>4</sup>, PhD; Lilian Schulze<sup>5</sup>, MSc; Barbara Posern<sup>5</sup>, MSc; Klemens Raile<sup>1</sup>, MD

<sup>1</sup>Charité – Universitätsmedizin Berlin, Department of Paediatric Endocrinology and Diabetes, Berlin, Germany

<sup>2</sup>Berlin Institute of Health, Berlin, Germany

<sup>3</sup>#dedoc° Diabetes Online Community, Dedoc Labs, Berlin, Germany

<sup>4</sup>Novo Nordisk Center for Biosustainability, Technical University of Denmark, Copenhagen, Denmark

<sup>5</sup>Designit Germany, Berlin, Germany

**Corresponding Author:**

Katarina Braune, MD

Charité – Universitätsmedizin Berlin

Department of Paediatric Endocrinology and Diabetes

Augustenburger Platz 1

Berlin, 13353

Germany

Phone: 49 30450616454

Email: [katarina.braune@charite.de](mailto:katarina.braune@charite.de)

## Abstract

**Background:** The COVID-19 pandemic poses new challenges to health care providers and the delivery of continuous care. Although many diabetes technologies, such as insulin pumps and continuous glucose monitors, have been established, the data from these devices are rarely assessed. Furthermore, telemedicine has not been sufficiently integrated into clinical workflows.

**Objective:** We sought to remotely support children with type 1 diabetes and their caregivers, enhance the clinical outcomes and quality of life of children with diabetes, increase multiple stakeholders' engagement with digital care via a participatory approach, evaluate the feasibility of using an interoperable open-source platform in a university hospital setting, and analyze the success factors and barriers of transitioning from conventional care to digital care.

**Methods:** Service design methods were used to adapt clinical workflows. Remote consultations were performed on a monthly and on-demand basis. Diabetes device data were uploaded from patients' homes to an open-source platform. Clinical and patient-reported outcomes were assessed before, during, and after the COVID-19 lockdown period in Germany.

**Results:** A total of 28 children with type 1 diabetes and their caregivers enrolled in this study and completed 6 months of remote visits. Of these 28 participants, 16 (57%) also opted to attend at least one of their regular visits remotely. After 3 months of remote visits, participants' time in range ( $P=.001$ ) and time in hyperglycemia ( $P=.004$ ) significantly improved, and their time in hypoglycemia did not increase. These improvements were maintained during the COVID-19 lockdown period (ie, between months 3 and 6 of this study). Participants' psychosocial health improved after 6 months.

**Conclusions:** Remote consultations and commonly shared data access can improve the clinical outcomes and quality of life of children with type 1 diabetes, even during challenging circumstances. A service design approach helped with the delivery of comprehensive and holistic solutions that accounted for the needs of multiple stakeholders. Our findings can inform the future integration of digital tools into clinical care during and beyond the pandemic.

**Trial Registration:** German Clinical Trials Register DRKS00016170; <https://tinyurl.com/skz4wdk5>

(*JMIR Mhealth Uhealth* 2021;9(4):e24374) doi:[10.2196/24374](https://doi.org/10.2196/24374)

**KEYWORDS**

telemedicine; telehealth; remote care; digital care; type 1 diabetes; pediatric diabetes; open source; service design; digital health; COVID-19; diabetes; workflow

## Introduction

People's interest in digital and remote care has been increasing worldwide. Although previous studies have demonstrated the effectiveness and increasing acceptability of telemedicine [1], knowledge on the implementation of digital care in health care settings and workflows remains limited.

The COVID-19 pandemic has forced many health care teams to find alternative approaches for delivering care to patients with chronic conditions. The demand for acute and emergency care has dramatically increased, and other areas of health care have been considerably compromised [2]. This is particularly concerning, as people with chronic conditions may have a higher risk for hospitalization and morbidity when they contract COVID-19. As such, health care professionals (HCPs) and public health institutions recommend that people who are at risk of SARS-CoV-2 infection should be protected from potential exposure to the virus. New legal frameworks that encourage health care teams to conduct remote, web-based consultations and prescribe medical software or apps have been introduced [3].

Diabetes is highly relevant to the field of telemedicine [4,5]. Intensive diabetes management has proven to be beneficial in delaying the onset of diabetes-related complications and reducing the severity of long-term complications [6]. Therapeutic guidelines have recommended a target hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>) level of <7.0% (ie, <53 mmol/mol) for people with type 1 diabetes [7,8]. However, many people with type 1 diabetes cannot meet these target HbA<sub>1c</sub> levels with standard care alone [9]. Modern treatment devices, such as insulin pumps and continuous glucose monitoring (CGM) sensors, are available to and widely used by people with diabetes in most industrialized countries. The uptake of insulin pumps and CGM sensors is highest in Western European countries [10-12] and the United States [9]. However, although less than a half of the population of people with type 1 diabetes in these regions use insulin pumps, the uptake of insulin pumps is much higher among children and adolescents. Germany, Switzerland, Luxembourg, and Austria have some of the highest technology uptake rates worldwide. For example, 92% of preschoolers in these countries used insulin pumps in 2017 [11]. The uptake of technology however decreases with age. Furthermore, although CGM sensors have been used by less than half of the population with diabetes (ie, typically young people with diabetes), CGM sensor uptake has increased considerably within the last 5 years in Germany and the United States [11,12]. However, to fully

benefit from these relatively modern treatment options, people with diabetes require a high level of self-management, diabetes care teams require expertise and training, and appropriate digital infrastructures must be available in health care settings.

Value-based and integrated care are promising strategies for managing chronic conditions, as they emphasize the importance of shared decision making and the organization of care. These strategies are largely information-driven processes. Technological tools for collecting and exchanging information are essential to all stakeholders involved in integrated care [13,14]. From a patient's perspective, integrated care aims to meet their health and social needs by using patient data as a starting point for redesigning their health care experiences.

Due to people's interest in the transition from traditional care to digital care, which has considerably increased as a result of the COVID-19 pandemic, evidence related to HCPs', patients', and caregivers' experiences with this transition has been emerging in various fields of medicine [15-28]. A variety of health care providers have shown increasing interest in the possibilities of digital health. However, there is little research on methods for integrating digital health tools into existing structures and workflows.

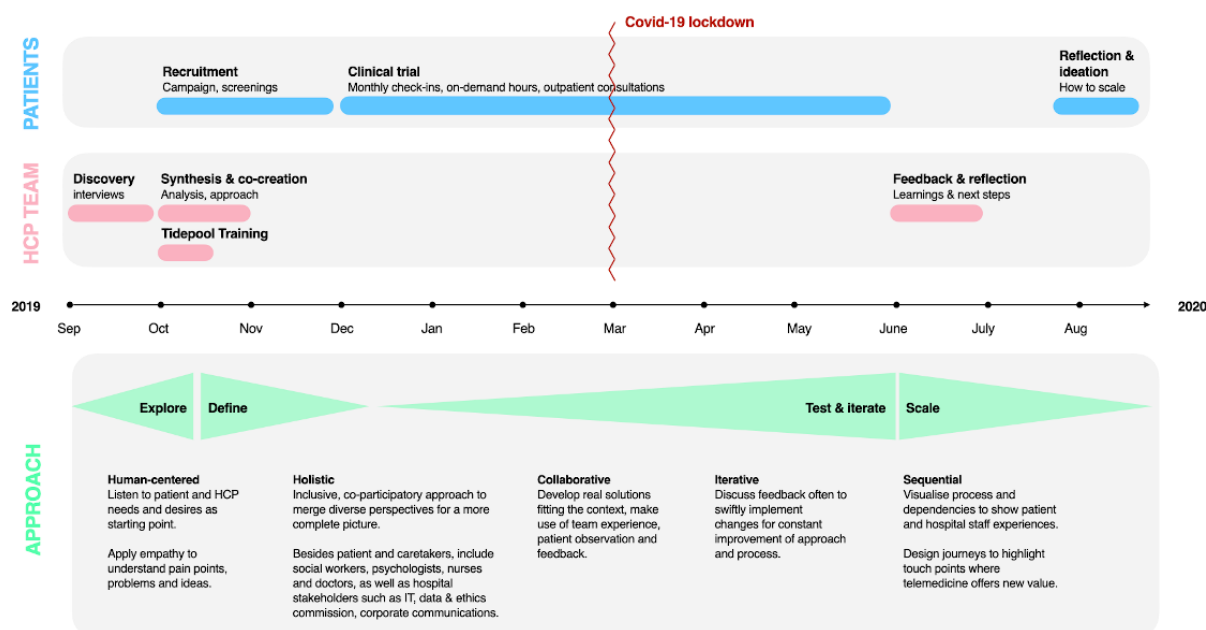
The Digital Diabetes Clinic (DDC) project sought to (1) remotely support children with type 1 diabetes and their caregivers during diabetes management; (2) increase the time that these children spend in the optimum glucose range (ie, time in range); (3) improve these children's quality of life; (4) increase multiple stakeholders' engagement with digital care via a participatory project that involves patients, caregivers, and care teams alike; (5) evaluate the feasibility of using an interoperable open-source platform to upload, store, and review diabetes device data in a university hospital setting; and (6) analyze the success factors and barriers of transitioning from conventional care to digital care.

## Methods

### Setting

This study was conducted in a tertiary, multidisciplinary pediatric diabetes care center of a university hospital. All participating HCPs were actively attending to children and adolescents with diabetes, and all participating families were receiving diabetes care from the care center. An overview of the methods that were used in this study and the clinical trial design is shown in [Figure 1](#).

**Figure 1.** An overview of the methods that were used in this study; the interventions that involved the HCP team, patients, and caregivers; and the service design thinking approach. This is presented on a project timeline. HCP: health care professional; IT: information technology.



## Service Design: A New Approach in Health Care

Service design is a strategic approach that is typically used in business environments and public sectors. This approach is used to create or improve processes for delivering desirable, consistent, and seamless experiences to users [29-31], such as staff, patients, and caregivers in health care settings. Service design thinking methods are used to assess the interrelations among different people, groups, organizations, resources, technologies, processes, and communication paths. This systemic and process-driven view is useful to design and deliver valuable services to users, including patients, their caregivers and health care professionals. Physical and digital touchpoints and resources are identified and evaluated, including data capture, usage and storage, as well as benefits and disadvantages of various communication channels (eg, apps, emails, phone calls, and written documentation).

### Workflow Analysis and Patients' Journeys

We followed the service design thinking process, starting with qualitative interviews. These interviews were conducted to assess users' experiences (ie, journeys) with receiving pediatric diabetes care from a university hospital and identify problem areas (ie, pain points), such as moments of frustration and challenges/gaps in the care process. The design team conducted semistructured interviews with the health care team (ie, physicians, diabetes educators, nurses, social workers, and psychologists) via an exploratory approach to learn about existing workflows, group-specific wishes, and stakeholders'

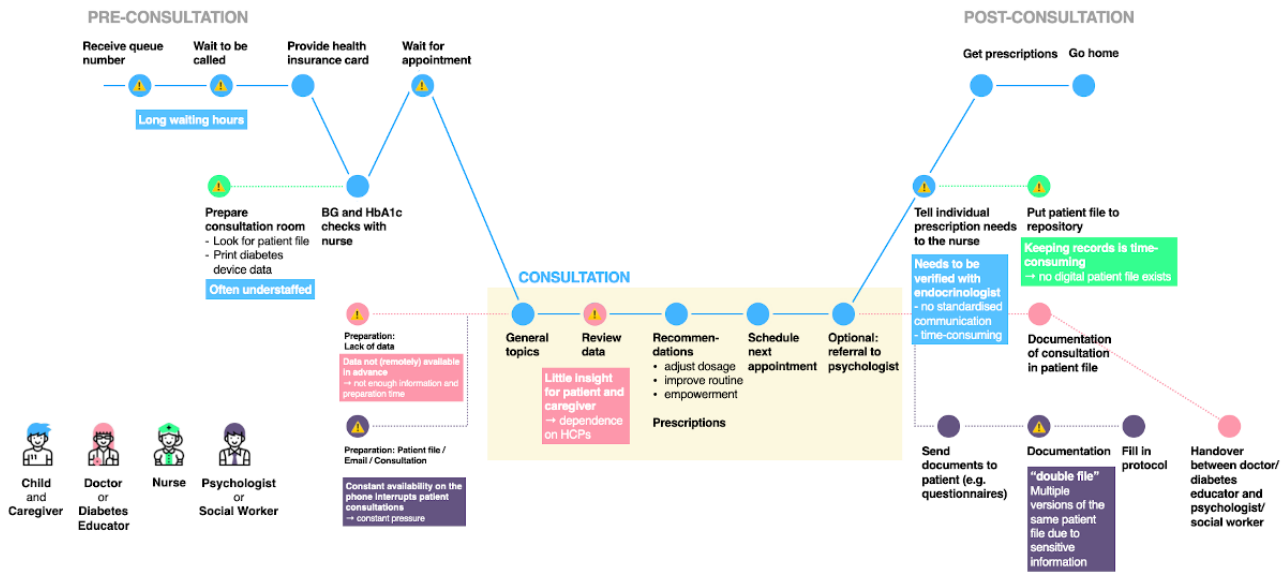
needs. The interviews were conducted on site (ie, at the hospital) and in person. Each interview involved a small group, which included 2 researchers and 2-4 interviewees. Voice recordings, notes, and photographs of the hospital workspace allowed us to better understand and portray technology use in the university hospital. Interviewees and interviewers used post-its and sketches to depict the process/steps that patients and caregivers undergo during hospital visits, such as scheduling appointments, waiting for consultations, waiting for laboratory test results, and receiving prescriptions. This information was used to map a comprehensive patient journey. The as-is journey is used to identify areas for improvements, by taking into account workflows and perspectives of various stakeholders (Figures 2-4).

Using findings from the as-is journey, a future journey was created that offers a better patient experience. Afterward, we developed a strategic concept with a clear value proposition, which showed how different stakeholders can better coordinate their activities so that users can access new or improved services [32,33].

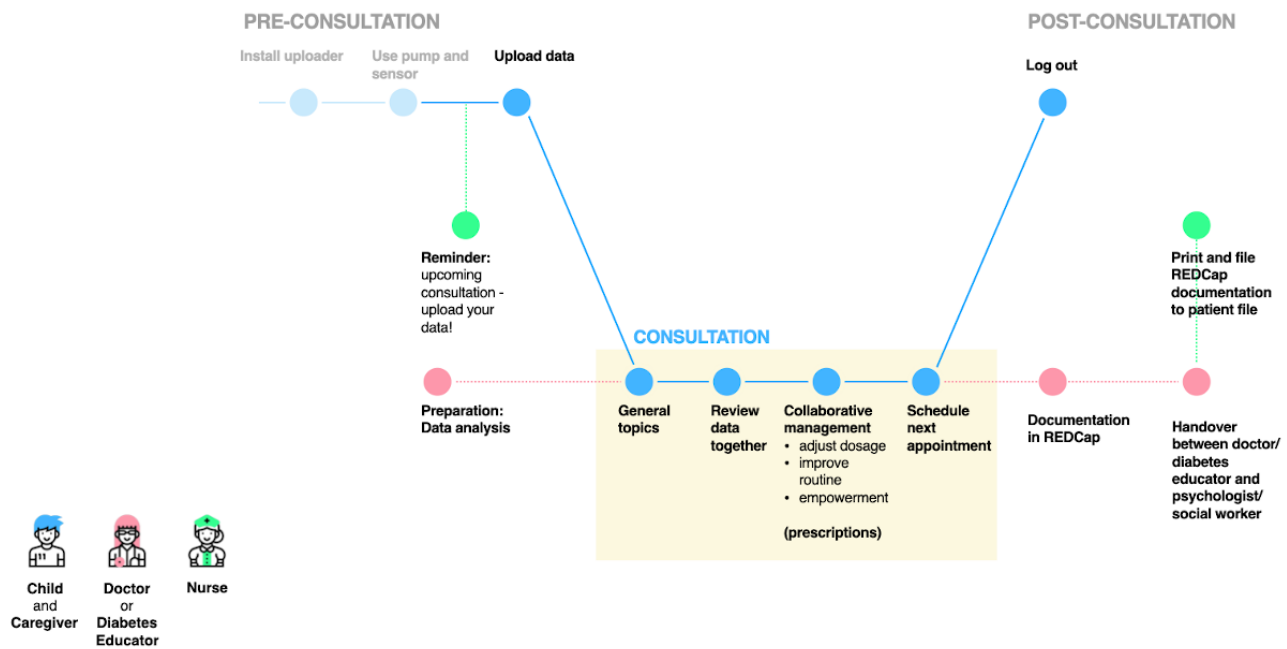
Once the concept was defined, we progressed to the implementation phase, during which the care team transitioned to their new responsibilities, adopted new technologies, and changed their routines and communication methods. This phase involved upskilling certain actors and transparently communicating these changes to patients, caregivers, care team members, organization members, or the general public.



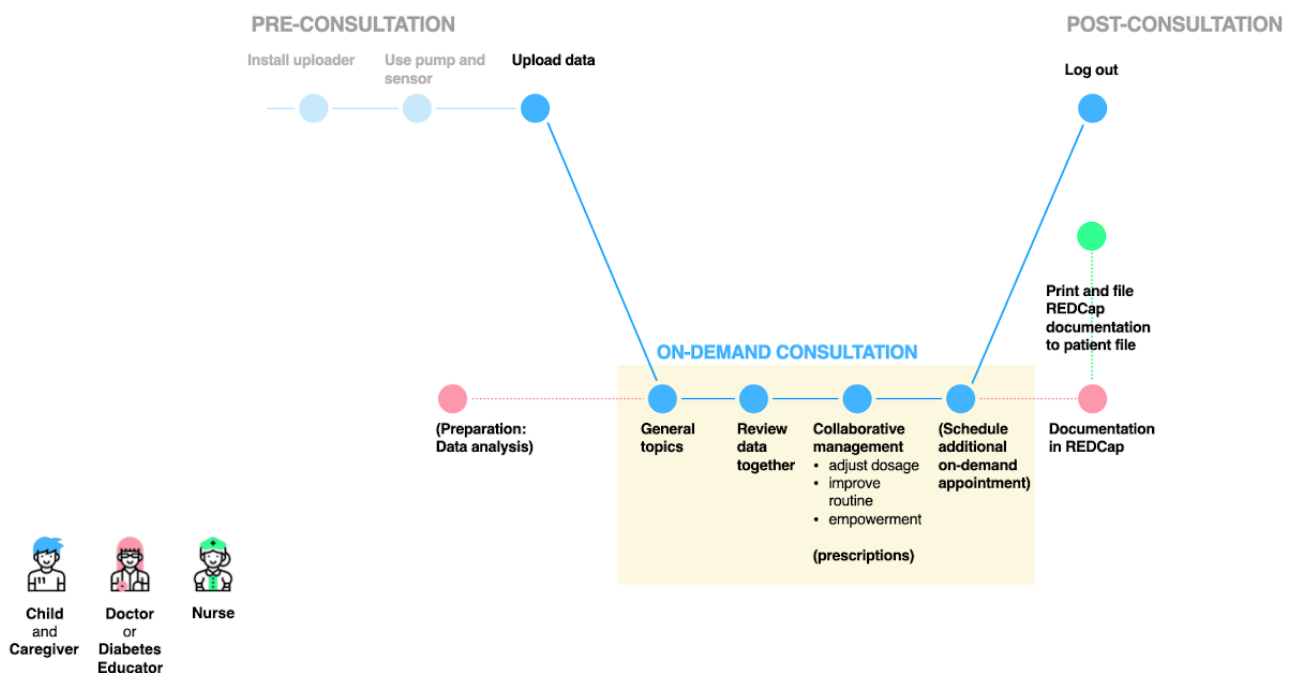
**Figure 2.** The "as-is" patient journey, the workflow of a conventional clinic visit, and current problem areas that require improvements. HbA<sub>1c</sub>: hemoglobin A<sub>1c</sub>; HCP: health care professional; BG: blood glucose.



**Figure 3.** The "future" patient journey and the workflow of remote monthly check-ins. REDCap: Research Electronic Data Capture.



**Figure 4.** The patient experience and the workflow of remote, on-demand consultations (ie, those conducted between monthly check-ins). REDCap: Research Electronic Data Capture.



## Stakeholder Workshops

A series of interactive workshops were held in preparation for and throughout the project. Initially, these workshops involved in-person meetings. Later, we transitioned to web-based meetings due to the COVID-19 pandemic. A participatory approach was used to facilitate a process for creating team trust; enabling clear communication; and promoting alignment, commitment, and shared goals. Relevant information about the project proposal was shared and discussed in order to identify concerns and expectations. Attendees were able to provide feedback and contribute to the design and organization of this study. The following questions were discussed: (1) “what do the terms ‘telemedicine’ and ‘digital care’ mean to us”; (2) “what are the pros and cons of a health care service that is carried out remotely”; (3) “how can health care settings, like a university hospital, adopt digital tools, new processes, and communication patterns into their current workflows”; (4) “which processes have to be adapted or redesigned”; (5) “what roles and responsibilities exist among the stakeholders”; and (6) “what technical and organizational support is needed”?

After the successful completion of 6 months of remote visits, a group of caregivers participated in a web-based workshop to share what they learned, discuss their reflections, and provide feedback. Similar to the previous workshops, our intention was to facilitate a participatory, cocreative session, during which caregivers could safely and comfortably voice opinions on how to improve users’ experiences with remote care. The study team created the following research questions: (1) “what was the study’s greatest benefit for you and your child”; (2) “if you could, what would be the one thing you would change”; (3) “what are the advantages or disadvantages of the DDC”; (4) “what elements should be adopted by existing health care services”; (5) “what challenges or benefits did you experience”;

(6) “how did the experience impact the children”; and (7) “did they give any feedback on the new process”?

## Clinical Trial Design

A prospective, longitudinal, open-label feasibility trial was conducted at a single clinical center from December 2019 to June 2020. Children with type 1 diabetes who were aged 3-12 years were enrolled. To be included in this trial, the child had to live with a caregiver who could upload data and attend remote visits. Moreover, the treatment at the time of the initiation visit had to include the use of an insulin pump for at least 6 months and the use of CGM sensor for at least 1 month. Children were asked for their consent to continue CGM during this study. Caregivers’ ability to use an insulin pump and CGM sensor was verified. This included their ability to insert the infusion set, change the reservoir or patch pump (ie, where applicable), calibrate CGM sensors, and read and interpret related data.

Families could not participate in the trial if the caregivers who operated the study tools were diagnosed with a physical or mental health condition that severely interfered with their ability to complete the study protocol. Families were also excluded if they had no access to a computer or if they were experiencing problems with their at-home cellular/Wi-Fi connection (ie, problems that interfered with their ability to upload data and attend video consultations).

## Recruitment

We made families aware of this study by placing posters in the hospital’s waiting area, mailing informational materials to all families with children that fulfilled the inclusion criteria, and directly contacting families’ pediatric endocrinologists. Additionally, a website was launched to disseminate information on this study’s overall goals, activities, technicalities, and recruitment process. Eligible participants were invited to workshops in small groups. During these workshops, families

learned about the study design and tools, provided their informed consent, and enrolled in this study.

### Hardware

Webcams with integrated microphones and speakers were provided to all participants. Caregivers could also use their own computers, tablets, or smartphone cameras. Access to a Windows personal computer or Macintosh computer with an internet connection was required to upload data.

### Data Platform

A personalized version of the open-source Tidepool platform (Tidepool Project) was used [34]. Caregivers could upload CGM sensor and insulin pump data to a secure and encrypted server that was hosted by Charité – Universitätsmedizin Berlin. They could also access their data via a web-based platform. As Tidepool is an open-source platform that is managed by a nonprofit group of parents of children with diabetes, its license permits third parties to use, reproduce, and alter the platform. Therefore, we were able to create a version of Tidepool that was adapted to the hospital's information technology infrastructure and local data protection requirements [35-38]. This decision was based on the platform's compatibility with multiple devices and manufacturers; our need to visualize CGM sensor and insulin pump data on a single platform in an integrated, device-agnostic fashion; the platform's easy upload process; the care team's and caregivers' need for equal data access at all times; the need for a user interface that was simple to understand; and the availability of ambulatory glucose profiles with information on all important clinical outcome parameters.

Caregivers signed up for an account, installed the Tidepool uploader on their computer, and provided data access to the care team. All participants received training on how to upload and review data. Several participants also used other software, such as Dexcom Clarity (Dexcom Inc) and Abbott FreeStyle LibreView (Abbott Diabetes Care Inc).

### Remote Visits

A secure, web-based video chat app (ie, Patientus [Jameda Gesellschaft mit beschränkter Haftung]) was used to conduct remote visits. As previously agreed upon in the stakeholder workshops, remote visits with a certified diabetes educator were scheduled on a monthly basis. This ensured that a frequent number of follow-ups were conducted and limited the additional burden on participating families at the same time. In addition, participants were offered optional, daily, on-demand consultation hours. Participants could also opt to remotely attend their routine appointments (ie, those that occurred every 2-3 months) with their pediatric endocrinologists. These appointments typically involved in-person visits, which were a part of participants' standard care.

### Outcome Measures

The primary outcome parameter was time in range, which refers to the percentage of time that participants spend with a sensor glucose level of 70-180 mg/dL (ie, 3.9-10.0 mmol/L).

Our secondary endpoints included the following: (1) time in hypoglycemia, which refers to the percentage of time that participants spend with a sensor glucose level of <54 mg/dL

(ie, 3.0 mmol/L) and 54-70 mg/dL (ie, 3.0-3.9 mmol/L); (2) time in hyperglycemia, which refers to the percentage of time that participants spend with a sensor glucose level of >250 mg/dL (ie, 13.9 mmol/L) and 180-250 mg/dL (ie, 10.0-13.9 mmol/L); (3) the incidence of severe hypoglycemia presenting with the need for assistance from others or unconsciousness; (4) the incidence and suspected cause of diabetic ketoacidosis; (5) diabetes-related hospitalizations; (6) HbA<sub>1c</sub> (ie, if available) and estimated HbA<sub>1c</sub> levels; (7) participants' quality of life, which parents reported via a web-based survey (ie, the Pediatric Quality of Life Inventory questionnaire); and (8) the feasibility of the care model, which was based on the proportion of participants who successfully complete 6 months of remote visits, as well as caregiver feedback from web-based surveys and interactive workshops (ie, optional feedback).

### Caregiver Feedback

At the end of the study, caregivers were invited to provide feedback via an optional web-based survey, which asked them to share the pros and cons of their experiences with remote care and their expectations and needs for future diabetes care.

### Data Collection and Analysis

A deidentified data set that included demographic data, outcome measures, and consultation information was documented and analyzed with the REDCap (Research Electronic Data Capture; Vanderbilt University) platform. REDCap is a secure electronic data capture tool that is hosted locally at Charité – Universitätsmedizin Berlin [39]. Furthermore, remote and in-person consultations were documented in the hospital's information system. Quantitative analyses were conducted with the R version 4.0.2 programming framework (The R Foundation), and the ggplot2 package was used to generate figures. Changes in primary and secondary outcome parameters were assessed with the Wilcoxon signed-rank test, which included a P value threshold of .05 for paired data and a 1-tailed test with an alternative hypothesis (ie, "less"). With regard to HbA<sub>1c</sub> descriptive statistics and associated statistical tests, missing HbA<sub>1c</sub> values were substituted with estimated HbA<sub>1c</sub> values (ie, when available).

### Ethical Conduct and Informed Consent

This study was conducted in accordance with the Declaration of Helsinki, Good Clinical Practice guidelines, and Data Privacy Law of Berlin (ie, the Berliner Datenschutzgesetz or Berlin Data Protection Act). This study was approved by the Charité ethics committee (approval number: EA2/125/18) and registered under the clinical trial registration number DRKS00016170. Informed consent was obtained from each child's caregiver prior to a family's inclusion in this study. A child-friendly version of the information sheet was provided to children aged 8-12 years, and they were asked for their assent to participate.

## Results

We present the results of a 6-month feasibility trial, information on the conceptual development of the trial, and details on the lessons that we learned throughout the process.

### Patient Journey and Workflow Analysis

The service design activities involved 5 pediatric endocrinologists, 3 diabetes educators, 3 nurses, 3 psychologists/social workers, 1 resident physician, and 4 families that previously participated in telemedicine interventions.

Figure 2 illustrates the patient journey and the current workflow of a conventional visit to the hospital. Multiple pain points were identified, such as long waiting hours; the lack of digitally available data from devices; patients'/caregivers' limited insight into their own data; and time-consuming tasks for the care team, such as documentation processes and the provision of paper file logistics. In the proposed workflow for scheduled remote consultations (Figure 3) and on-demand remote consultations (Figure 4), the identified problem areas were addressed; HCPs and patients/caregivers were provided with equal access to data, and they collaboratively agreed upon individual therapy goals and methods for achieving them. The documentation process was simplified.

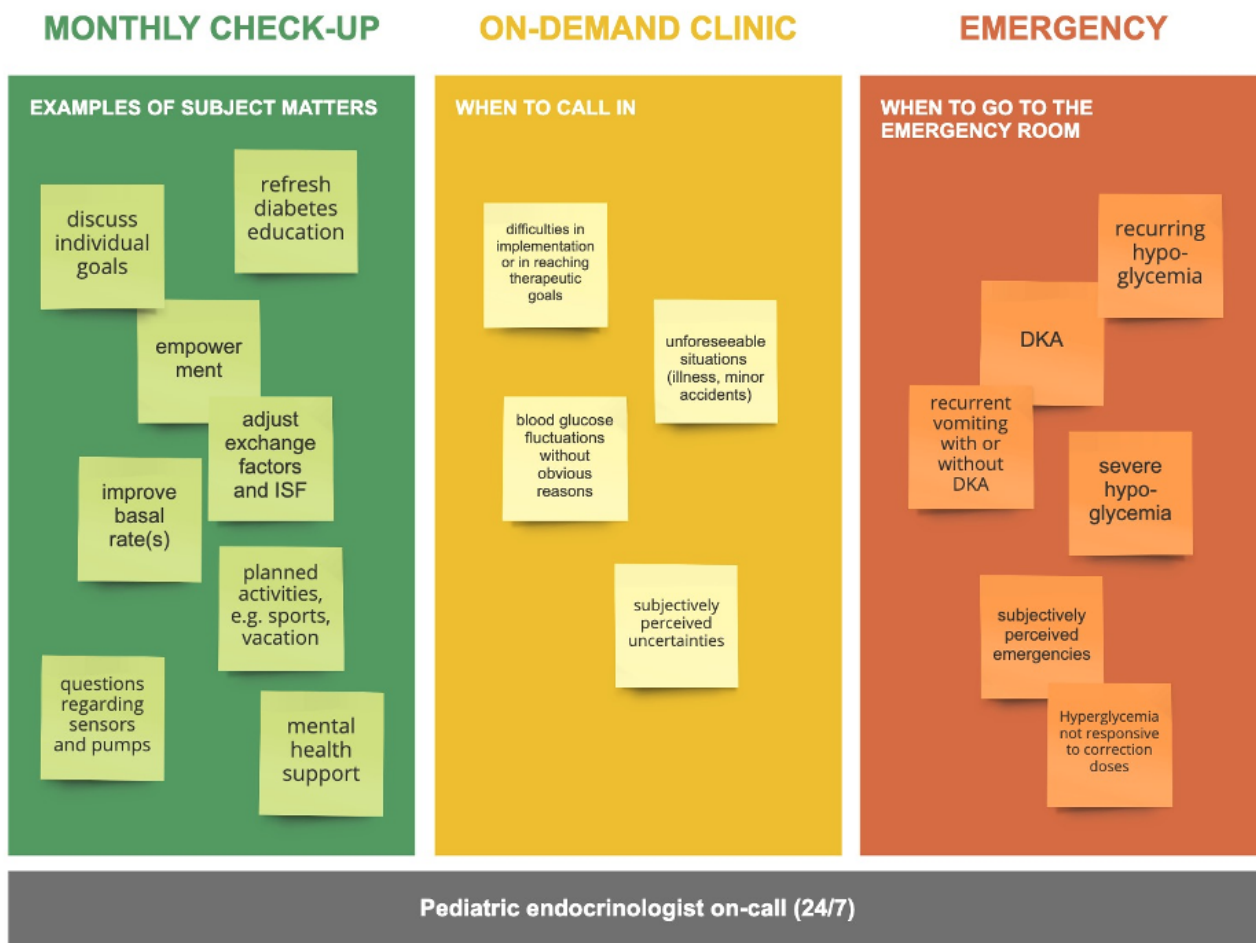
### Stakeholder Workshops

The stakeholders agreed that the project's value proposition was to improve patients' well-being by using existing ambulatory services to complement remote consultations instead

of completely replacing ambulatory services, and by improving data access and analysis. Much time was spent on structuring a remote service that promotes better interactions, higher flexibility, and lower stress among all stakeholders. The clear advantages of telemedicine were identified, such as the ability to save time and effort and the possibility of delivering health care to families in rural areas. Furthermore, people believed that digital care was less bureaucratic than nondigital care. Hence, digital care was less time-consuming than nondigital care, as per the documentation. Therefore, health care providers could save money by using digital care services. The care team was also generally open to the adoption of new technology, as they saw clear benefits for patients. Limited staff availability due to the increasing economization of the health care sector, organizational challenges, and structural challenges were perceived to be the main barriers to the adoption of new tools and care pathways. Furthermore, safety concerns were addressed and a triage system for emergency scenarios was created (Figure 5).

The perceived challenges that were identified in this study were mainly technical in nature. Such challenges included the compatibility between several sensors/pumps and the Tidepool uploader and hospital server errors that required time-consuming and complex solutions.

**Figure 5.** The triage system for on-demand consultations; risk assessments; and the emergency management of severe hypoglycemia, DKA, or other emergencies that are subjectively perceived as serious. DKA: diabetic ketoacidosis; ISF: insulin sensitivity factor.



### Study Cohort

A total of 28 patients (age: median 8 years, SD 2.6 years) with diabetes (duration: median 4 years, SD 2.2 years) were enrolled in this study. It should be noted that 3 more families were interested in participating in this study, but they could not be included due to the incompatibility of their diabetes hardware (n=1), their lack of access to a computer (n=1), and technical issues with their at-home personal computers (n=1). Of the 27 participating families, 1 (4%) had 2 children with diabetes, 2 (7%) came from single-parent households, 2 (7%) had parents who lived separately, and 7 (26%) had a migration background (ie, at least 1 parent was born in a country other than their

country of residence). All 28 children used insulin pumps and CGM sensors as a sensor-augmented pump therapy system (n=27) or hybrid closed-loop system (n=1). Families had a median annual household net income of €60,000 (US \$72,824.10; SD €34,078 [US \$41,361.70]), and 45% (28/62) of caregivers had a university degree. The income and education levels of the cohort were above the national average. Children's CGM sensor and pump supplies were fully covered by their health insurance; 85% (23/27) of the families had a public health care plan and 15% (4/27) had a private health care plan. A summary of the cohort's demographic characteristics are reported in [Table 1](#).

**Table 1.** Sociodemographic characteristics of the study cohort.

Characteristics	Value
<b>Gender, n (%)</b>	
Female	18 (64.2)
Male	10 (35.7)
Other	0 (0)
<b>Comorbidities, n (%)</b>	
Celiac disease	3 (10.7)
Hashimoto thyroiditis	1 (3.6)
Other	2 (7.1)
<b>Type of insulin pump, n (%)</b>	
Medtronic 670G	1 (3.6)
Medtronic 640G	23 (82.1)
Medtronic Veo	4 (14.3)
<b>Type of continuous glucose monitoring sensor, n (%)</b>	
Medtronic Guardian	13 (46.4)
Dexcom G6	5 (17.9)
Dexcom G5	1 (3.6)
FreeStyle Libre 2	9 (32.1)
<b>Caregiver's/mother's employment status, n (%)</b>	
Full-time employment	10 (35.7)
Part-time employment	13 (46.4)
Unemployed	4 (14.3)
Student	1 (3.6)
Not available	0 (0)
<b>Caregiver's/mother's highest educational level, n (%)</b>	
No/some high school	1 (3.6)
High school	4 (14.3)
University degree or diploma	12 (42.9)
Doctorate	2 (7.1)
Other	9 (32.1)
Not available	0 (0)
<b>Caregiver's/mother's professional background, n (%)</b>	
Health care/science	8 (28.6)
Education and childcare	3 (10.7)
Information technology	1 (3.6)
Service	3 (10.7)
Other	9 (32.1)
None	1 (3.6)
Not available	0 (0)
<b>Caregiver's/father's employment status, n (%)</b>	
Full-time employment	22 (78.6)
Part-time employment	2 (7.1)
Unemployed	2 (7.1)

Characteristics	Value
Student	1 (3.6)
Not available	1 (3.6)
<b>Caregiver's/father's highest education level, n (%)</b>	
No/some high school	4 (14.3)
High school	3 (10.7)
University degree or diploma	13 (46.4)
Doctorate	1 (3.6)
Other	6 (21.4)
Not available	1 (3.6)
<b>Caregiver's/father's professional background, n (%)</b>	
Health care/science	2 (7.1)
Education and childcare	1 (3.6)
Information technology	4 (14.3)
Service	2 (7.1)
Other	14 (50)
None	2 (7.1)
Not available	1 (3.6)

## Feasibility

All enrolled participants completed 6 months of monthly remote visits. In addition, 57% (16/28) of the participants opted to remotely attend at least one of their regular clinic visits (ie, those that occurred every 2-3 months) with their pediatric endocrinologist. The on-demand clinic service was used by 29% (8/27) of the families. Of these 8 families, 7 (88%) made use of this service once, and 1 (12%) used this service multiple times throughout the study. The subject matters that were discussed during remote visits are summarized in Table 2. In 96.4% (118/122) of the consultations, participants felt confident

with remotely uploading, accessing, and reviewing their data. Although the technical aspects of the data and video chat platforms were mostly discussed during the first web-based visit, follow-up visits largely focused on diabetes- and health-related topics. During the monthly check-ups, 90.4% (110/122) of children fully achieved their individual therapy goals, and 6.1% (7/122) partially achieved their individual therapy goals. There were no study dropouts. Severe hypoglycemia, diabetic ketoacidosis, or other issues that required study personnel to consult the on-call endocrinologist did not occur. Patient handovers to psychologists and social workers occurred 6 times.

**Table 2.** Activities and subject matters that were discussed during remote care visits.

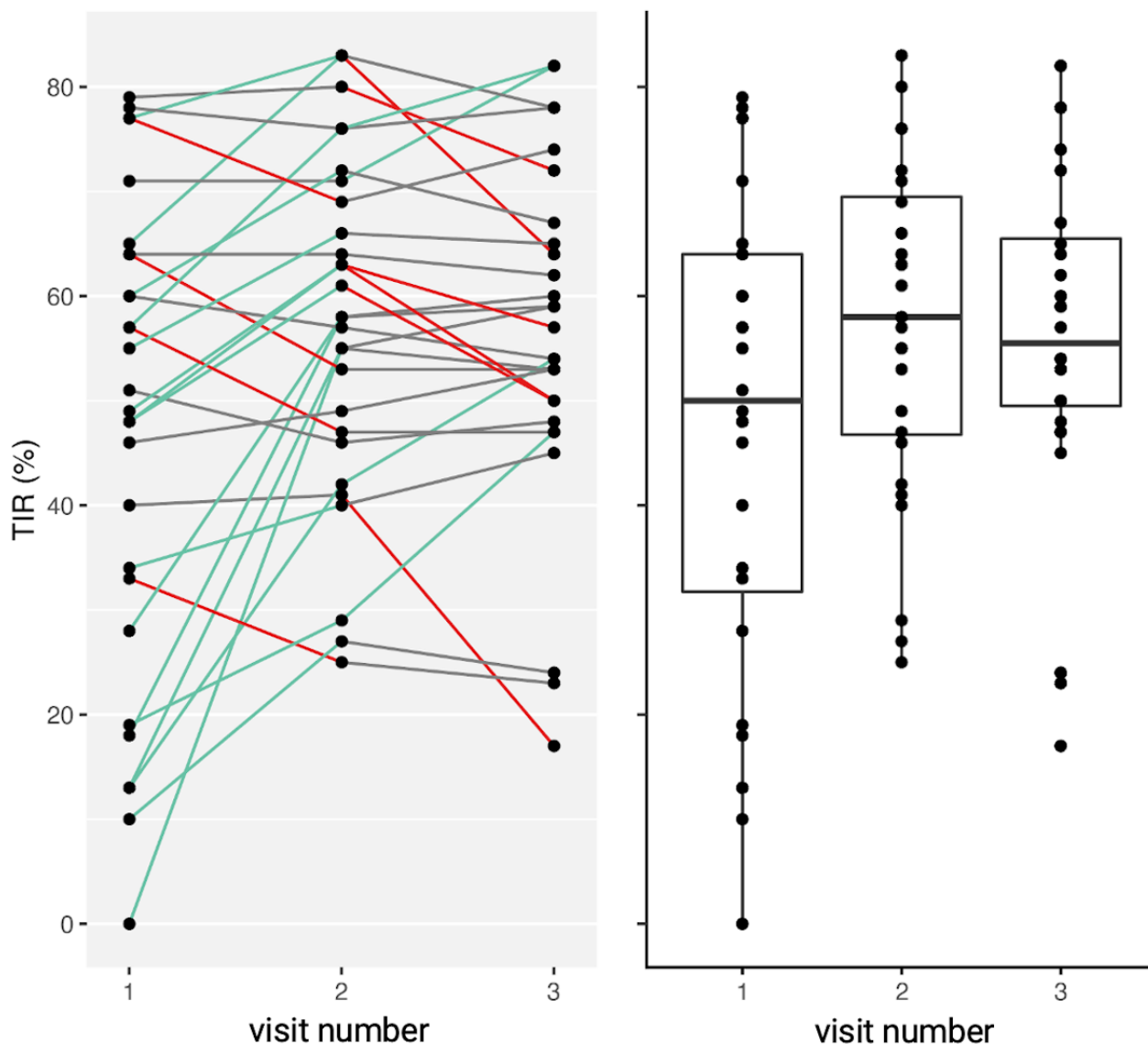
Activities and subject matters	Monthly check-in, n (%)	On-demand clinic, n (%)
Reviewed data together	96 (85.7)	6 (60)
Basal rate adjustments	59 (52.7)	4 (40)
Carbohydrate exchange factor adjustments	52 (46.4)	3 (30)
Technical aspects of the data platform	34 (30.4)	1 (10)
General organizational matters	32 (28.6)	1 (10)
Refreshed diabetes education and training	25 (22.3)	2 (20)
Technical aspects of the video chat platform	18 (16.1)	2 (20)
Mental health concerns	13 (11.6)	1 (10)
Technical aspects of continuous glucose monitoring	9 (8)	1 (10)
Technical aspects of insulin pumps	8 (7.1)	1 (10)
Therapy adjustments for physical activity	7 (6.3)	0 (0)
Insulin sensitivity factor adjustment	5 (4.5)	1 (10)
Acute illness	2 (1.8)	1 (10)
Other topics	30 (26.8)	3 (30)

### Clinical and Patient-Reported Outcomes

After completing 3 months of remote consultations, participants' time in range ( $P=.001$ ) and time in hyperglycemia ( $P=.004$ ) significantly improved (Figure 6). This improvement was maintained over the lockdown period (ie, between months 3 and 6 of this study); daycares, playgrounds, schools, universities, nonessential businesses, and international borders were closed

during the first wave of the COVID-19 pandemic in Germany (ie, from March to May 2020). After 3 and 6 months of remote visits, patients' time in hypoglycemia did not significantly increase (3 months:  $P=.21$ ; 6 months:  $P=.08$ ), no significant changes in HbA<sub>1c</sub> levels were observed (3 months:  $P=.43$ ; 6 months:  $P=.42$ ), and patients' psychosocial health significantly improved. All details on outcome parameters are shown in Table 3.

**Figure 6.** The individual changes and overall distribution of participants' percent TIR of sensor glucose targets before remote consultations (ie, visit 1), after 3 months of remote consultations (ie, visit 2), and after 6 months of remote consultations (ie, visit 3). Green lines indicate that the individual change in TIR was  $>5\%$  (ie, high amount of change). Red lines indicate that the individual change in TIR was  $<5\%$  (ie, low amount of change). Grey lines indicate that the individual change in TIR was within 5% of the acceptable range (ie, a stable TIR). TIR: time in range.





**Table 3.** Clinical and patient-reported outcomes before, during, and after 6 months of remote consultations. All significance levels for the visit 2 and visit 3 results were compared to those for visit 1.

Outcome measures	Visit 1 (ie, enrollment)		Visit 2 (ie, 3 months of completed remote consultations)		Visit 3 (ie, 6 months of completed remote consultations)	
	Value, mean (SD)	P value	Value, mean (SD)	P value	Value, mean (SD)	P value
Time in range, %	46.9 (22.9)	Referent	57.5 (16.3)	.001	56.3 (16.5)	.02
<b>Time in hypoglycemia (mg/dL [mmol/L]), %</b>	5.1 (12.1)	Referent	4.7 (6.9)	.21	3.7 (4.3)	.08
54-70 (3.0-3.9)	3.1 (5.4)	Referent	3.5 (5.0)	.17	2.7 (3.0)	.11
<54 (3.0)	2.0 (6.9)	Referent	1.2 (2.1)	.39	1.1 (1.4)	.12
<b>Time in hyperglycemia (mg/dL [mmol/L]), %</b>	48.0 (26.2)	Referent	37.9 (18.8)	.004	40.0 (18.7)	.02
180-250 (10.0-13.9)	26.7 (14.4)	Referent	25.9 (13.2)	.35	26.1 (10.0)	.47
>250 (13.9)	21.3 (20.2)	Referent	12.0 (12.8)	.004	13.9 (14.7)	.01
Hemoglobin A <sub>1c</sub> level, %	7.5 (0.9)	Referent	7.7 (0.9) <sup>a</sup>	.43	7.6 (1.1)	.42
Hemoglobin A <sub>1c</sub> level, mmol/mol	58.2 (9.0)	Referent	58.1 (9.3) <sup>a</sup>	.43	59.1 (13.0)	.42
<b>Quality of life score<sup>b</sup></b>						
Psychosocial health summary score	72.5 (14.0)	Referent	— <sup>c</sup>	—	78.0 (14.4)	.04
Physical health summary score	81.0 (17.9)	Referent	—	—	82.7 (16.0)	.14
Diabetes-related score	70.3 (10.7)	Referent	—	—	73.6 (9.0)	.06

<sup>a</sup>Not available for patients who opted to perform their month-3 visit remotely. Instead, estimated hemoglobin A<sub>1c</sub> values were calculated based on sensor data.

<sup>b</sup>Based on Pediatric Quality of Life Inventory scores.

<sup>c</sup>Not available.

## Caregiver Feedback

Caregivers expressed that remote care was beneficial to them and their child. They also believed that remote care had advantages over in-person meetings. Reductions in the amount of time and stress (ie, those associated with hospital visits); flexibility during different times of the day; and the opportunity to be in a safe, comfortable, and familiar setting allowed for more engagement and dedicated interactions between families and health service providers. Families also believed that health service providers were less stressed, more dedicated, and focused. The ability to review data together resulted in new and valuable insights for most stakeholders and enabled caregivers to take initiative and make adjustments to therapy. The families were satisfied with the care that they received. They also expressed their desire to continue digital care after their participation in this study and suggested that digital care should be fully integrated into routine care. Furthermore, families desired remote consultations with psychologists and social workers. The intervals between visits were shorter, and caregivers perceived this as beneficial for discussing any arising questions. The late afternoon and evening hours were caregivers' preferred times for attending consultations. Technical problems with internet connections, video chats, and data platforms occasionally occurred.

Finally, remote care was perceived to be "more modern, timely and suited to their needs," with a perceived general improvement of the children's well-being, improved glycemic outcomes, and

newly gained insights. Details on caregiver feedback are shown in Table S1 in [Multimedia Appendix 1](#).

## Discussion

### Principal Results

This study describes the feasibility of remote care, the transition from traditional care to digital care, and the use of an interoperable health data platform in an ambulatory pediatric diabetes care setting. Digital care was delivered successfully, and participants were satisfied with the care that they received. Remote and continuous data access considerably improved for patients, caregivers, and the health care team. Data access was perceived as helpful for therapeutic decision making. Despite the COVID-19 pandemic and its numerous potential implications for physical and mental health [40,41] and people's social and family lives, better clinical outcomes were achieved before the intervention period, and these outcomes were maintained during the intervention period. Furthermore, patients' psychosocial health significantly improved.

Our findings are in line with those of studies that were conducted before [4,5,42-44] and during [45-51] the COVID-19 pandemic (ie, studies that reported on the benefits of remote care in various regions and settings). Our study is the first to report on the following: (1) the impact that remote care during the COVID-19 pandemic has on clinical outcomes; and (2) the integration of remote care into pediatric diabetes care during the COVID-19 pandemic. Our study is also the first to describe how service

design methods were used to increase stakeholder engagement and improve workflows in clinical care.

Conventional hospital visits and admissions provide patients with artificial time windows that rarely address their actual daily challenges and individual needs. Young people with diabetes and caregivers have often perceived a lack of communication with their health care providers when it comes to discussing their treatments [52]. They have also felt that their care teams do not sufficiently understand their daily problems and do not have a lot of involvement in the therapeutic decision-making process [52]. Studies have shown that few people with diabetes and caregivers regularly download and review their own data [53,54]. It has also been reported that self-assessments and specialists' referrals (ie, those for obtaining data) can improve diabetes management [55].

The findings from this study allowed us to validate the technical instruments and organizational processes that are necessary for the transition to digital care and new workflows. Both the "as-is" journeys and "future" journeys were key tools for promoting shared understanding and alignment among stakeholders. As recent modalities have led to overall changes in work culture and social interactions (eg, the majority of professional meetings, social gatherings, university coursework, and school classes have been conducted via web-based platforms), there is a strong need for the health care sector to adopt new methods for communication. The implementation of remote data assessments and web-based communications can therefore improve people's access to specialty care during and beyond the pandemic. However, a paradigm shift in the delivery, management, and funding of health care services is required.

The use of an open-source data platform was considered positive for the following reasons: (1) it was interoperable with devices from different manufacturers; (2) it allowed both parties to immediately access data; and (3) it provided a positive user experience. However, technical issues occasionally occurred throughout the intervention period. These issues were sometimes challenging to solve, as the data platform was not a plug-and-play service that was provided by a third party. Therefore, follow-up projects should include structured management plans that clarify the roles and responsibilities of technical support personnel.

During the transition to technological innovations, the digital divide might leave several user groups behind. Therefore, while people with diabetes have generally reported positive experiences with diabetes technology, the complexity of accessing and maintaining such technology remains a challenge. It should be noted that younger adults, who are generally perceived as "tech savvy," are in fact less likely to embrace the use of diabetes technology. Additionally, physical barriers (eg, the need to carry devices) and general diabetes distress are more severe among younger adults than older adults [56].

In this study, the team primarily interacted with caregivers. Further research is needed on actively increasing children's and adolescents' involvement in telemedicine consultations and adapting digital care interventions to cater to the needs of other age groups, people with diabetes (ie, other than type 1 diabetes), and people who use self-monitoring blood glucose devices and

metered dose inhalers (eg, uploading data from glucometers and electronic pens or integrating diabetes diary apps into routine care). Based on the stakeholder feedback, we conclude that it is crucial to integrate new digital tools into routine care instead of creating separate care pathways. Therefore, further service design research should address how health care teams in other health care settings (eg, high-volume clinics for adult diabetes care) and staff with limited technical skills can be trained to confidently provide digital care.

The ability to effectively provide telemedicine increases if a clinic has experience with diabetes technology provision (eg, insulin pumps, CGM sensors, electronic health records, diabetes management software, etc) and the required infrastructure (eg, software, computers, and internet connections). However, it is worth acknowledging that although the use of diabetes technology is clearly associated with countries' local reimbursement strategies or health insurance plans [57], diabetes technology uptake is heterogeneous in countries that offer full reimbursements. Individual aspects such as personal attitudes or interests (ie, those of people with diabetes and, more importantly, HCPs), awareness, structure, and capacity are insurance-related determinants to improving the accessibility of diabetes technology [58]. Thus, we are aware that our proposed model may not be suitable for all pediatric diabetes clinics.

The beliefs and attitudes of HCPs may be barriers to increasing the universal accessibility of advanced diabetes technologies [56,59,60]. New care models may initially result in feelings of uncertainty, and they might seem overwhelming to people who have usually delivered in-person care for the past few decades [61]. Organizational and structural changes can lead to frustration and negatively impact people's motivations for adopting new care pathways. To ensure that health care teams do not shy away from new technology and additional work requirements, all relevant stakeholders must be engaged with the transition process as early as possible. The business environment has learned that the service design approach can be used to innovate methods for addressing people's needs. Therefore, the health care ecosystem can greatly benefit from using the same approach [31]. As an applied research and innovation framework, service design prioritizes empathy for the users of a service or product; embraces interdisciplinarity and collaboration within project teams; and encourages the action-oriented, rapid prototyping of user-derived insights instead of top-down hypotheses. Service design has proven to be beneficial for encouraging all stakeholders to contribute their ideas during the design process, acknowledge their concerns, and build supportive practices [61].

Our study adds to the ongoing discussion on the importance of time in range (ie, as a measure that is comparable to or more important than HbA<sub>1c</sub> levels). Our choice to use time in range as the primary outcome parameter and HbA<sub>1c</sub> level as a secondary outcome parameter may be a strength of our study (ie, compared to most other studies that have reported on the clinical outcomes of people with diabetes). Although our study cohort's time in range and quality of life significantly improved, there were no significant changes in patients' HbA<sub>1c</sub> levels.

Although HbA<sub>1c</sub> is widely used as a primary outcome parameter in many CGM studies [62-66], time in range has been acknowledged as an outcome measure for representing glycemic control [67,68]. Studies have only shown moderate correlations between time in range and associated HbA<sub>1c</sub> levels (ie, changes in HbA<sub>1c</sub> levels for a given change in time in range widely varies) [69]. Our findings support the fact that there is a need to use CGM-derived outcome measures in clinical care and clinical research and a need to identify new surrogate markers for the development of diabetes-related long-term complications [67-69]. These needs can be addressed, as CGM data can be remotely assessed without having to resort to additional invasive procedures for people with diabetes. Furthermore, clinical outcome improvements may remain unobserved if researchers only focus on analyzing HbA<sub>1c</sub> levels.

We acknowledge that our study has several strengths and limitations. First, this study provides other diabetes care teams with a practical example of how to take advantage of the opportunities that have arisen from the necessity of remote care. Second, our study supports the use of technology in the delivery of diabetes care and the promotion of patient involvement in the cocreation of services. Third, our study is based on real-world data; it presents the different perspectives of health service providers and users. The limitations of this study should also be acknowledged. This study was a single-arm, nonrandomized feasibility trial that analyzed data from a small cohort of patients. Due to the trial's observational nature and our lack of a control group, our ability to assess the effectiveness of remote and routine care for all people with diabetes was limited. Although our participants had widely varying characteristics (ie, various education levels, income levels, and professional backgrounds), the majority of the participants were

from middle- to high-income and educated families, and all participants used insulin pumps and CGM sensors. This may mean that socioeconomic status-related biases are present in our study. Furthermore, people who experience language barriers and people with low levels of technological literacy might not have felt confident with participating in this study. This might indicate that selection bias was present in our study. Additionally, access to the internet and a computer was required. Although these technologies are available to most families, they were not available to all eligible families. Furthermore, as CGM sensors, insulin pumps, and supplies are fully covered by the public and private health insurance plans in Germany (ie, for children with diabetes), this project may not be applicable to all diabetes care settings. Limited access to diabetes technologies, which is evident in many regions outside of Western Europe, could limit the applicability of our service design approach. Although advances in technologically mediated treatments are promising, there are still concerns about social inequality and the challenge of ensuring that such treatments are widely disseminated across the population. More research is needed to understand these potential obstacles and provide appropriate education and support.

### Conclusions

This study sought to identify and solve the following problems in diabetes care: the limited accessibility of diabetes device data; the poor interoperability of data from different devices; and restricted access to specialists, especially during a global pandemic. Our study design allowed the care team, patients, and caregivers to actively contribute to the DDC project and promoted shared decision making. The results generated by this study will help to inform and improve methods for implementing remote and digital diabetes care into the wider health care sector during and beyond the pandemic.

---

### Acknowledgments

We would like to thank all children with diabetes, their families, and all professional diabetes team members from the Interdisciplinary Social Pediatric Center and the Department of Paediatric Endocrinology and Diabetes of Charité – Universitätsmedizin Berlin, who greatly contributed to this study. Furthermore, we would like to thank the team members of Tidepool, Diabeloop, and Sensotrend for their advice on implementing the Tidepool platform and uploader in our study setting. We also acknowledge the German Research Foundation and the Open Access Publication Funds of Charité – Universitätsmedizin Berlin for their support with the publication of this study.

The DDC project was funded by the Berlin Institute of Health. The project also received nonfinancial support from the service design company Designit. The funding source was not involved with this study.

---

### Authors' Contributions

KBraune performed the literature search. KBraune, KBoss and KR designed the study. KBraune, KBoss, JSH, LS, and BP collected the data. KBraune, AT, LS, and BP analyzed and interpreted the data. KBraune wrote the initial draft of the manuscript. All authors critically reviewed and revised the draft. All authors read and approved the final version of the manuscript. KBraune is the guarantor of this study.

---

### Conflicts of Interest

All authors completed the Unified Competing Interest form. KBraune received research grants from the Berlin Institute of Health Junior Clinician Scientist program for conducting this study. Outside of this study, KBraune received research grants from the Berlin Institute of Health Digital Clinician Scientist program, the European Commission's Horizon 2020 Research and Innovation program, Wellcome Trust, Stiftung Charité, and the German Diabetes Association. Outside of this study, KBraune served as a

speaker and advisory board member for Medtronic Diabetes and received fees for medical consulting and public speaking from Roche Diabetes Care, Dexcom, Medtronic Diabetes, Diabeloop, and Bertelsmann Stiftung. JSH is employed as a part-time diabetes coach at mySugr. KR received grants from the European Commission's Horizon 2020 Research and Innovation Program (grant agreement number: 823902). Outside of this study, KR is an advisory board member of Lilly Diabetes Care and Abbott Diabetes Care. All other coauthors have no conflicts of interest to declare.

## Multimedia Appendix 1

Table S1. Caregiver feedback regarding their experience with remote consultations.

[[DOCX File, 19 KB - mhealth\\_v9i4e24374\\_app1.docx](#)]

## References

1. Bashshur RL, Shannon G, Krupinski EA, Grigsby J. Sustaining and realizing the promise of telemedicine. *Telemed J E Health* 2013 May;19(5):339-345 [FREE Full text] [doi: [10.1089/tmj.2012.0282](https://doi.org/10.1089/tmj.2012.0282)] [Medline: [23289907](https://pubmed.ncbi.nlm.nih.gov/23289907/)]
2. Pfiester E, Gomber A, Rojas D, Al-Bajar M, Wofford K, Ortiz L, et al. Implications of the covid-19 pandemic on chronic care: global perspectives from people with type 1 diabetes. *BMJ Opinions*. 2020 Nov 06. URL: <https://blogs.bmj.com/bmj/2020/11/06/implications-of-the-covid-19-pandemic-on-chronic-care-global-perspectives-from-people-with-type-1-diabetes/> [accessed 2021-02-26]
3. Entwurf eines Gesetzes für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG). *Digitales Versorgungsgesetz*. URL: [https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3\\_Downloads/Gesetze\\_und\\_Verordnungen/GuV/D/Digitale-Versorgung-Gesetz\\_DVG\\_Kabinett.pdf](https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/D/Digitale-Versorgung-Gesetz_DVG_Kabinett.pdf) [accessed 2020-09-04]
4. Levin K, Madsen JR, Petersen I, Wanscher CE, Hangaard J. Telemedicine diabetes consultations are cost-effective, and effects on essential diabetes treatment parameters are similar to conventional treatment: 7-year results from the Svendborg Telemedicine Diabetes Project. *J Diabetes Sci Technol* 2013 May 01;7(3):587-595 [FREE Full text] [doi: [10.1177/193229681300700302](https://doi.org/10.1177/193229681300700302)] [Medline: [23759390](https://pubmed.ncbi.nlm.nih.gov/23759390/)]
5. Wood CL, Clements SA, McFann K, Slover R, Thomas JF, Wadwa RP. Use of telemedicine to improve adherence to american diabetes association standards in pediatric type 1 diabetes. *Diabetes Technol Ther* 2016 Jan;18(1):7-14. [doi: [10.1089/dia.2015.0123](https://doi.org/10.1089/dia.2015.0123)] [Medline: [26295939](https://pubmed.ncbi.nlm.nih.gov/26295939/)]
6. Diabetes Control and Complications Trial Research Group, Nathan DM, Genuth S, Lachin J, Cleary P, Crofford O, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993 Sep 30;329(14):977-986. [doi: [10.1056/NEJM199309303291401](https://doi.org/10.1056/NEJM199309303291401)] [Medline: [8366922](https://pubmed.ncbi.nlm.nih.gov/8366922/)]
7. American Diabetes Association. 6. Glycemic targets: Standards of medical care in diabetes-2018. *Diabetes Care* 2018 Jan;41(Suppl 1):S55-S64. [doi: [10.2337/dc18-S006](https://doi.org/10.2337/dc18-S006)] [Medline: [29222377](https://pubmed.ncbi.nlm.nih.gov/29222377/)]
8. DiMeglio LA, Acerini CL, Codner E, Craig ME, Hofer SE, Pillay K, et al. ISPAD Clinical Practice Consensus Guidelines 2018: Glycemic control targets and glucose monitoring for children, adolescents, and young adults with diabetes. *Pediatr Diabetes* 2018 Oct;19 Suppl 27:105-114. [doi: [10.1111/pedi.12737](https://doi.org/10.1111/pedi.12737)] [Medline: [30058221](https://pubmed.ncbi.nlm.nih.gov/30058221/)]
9. Foster NC, Beck RW, Miller KM, Clements MA, Rickels MR, DiMeglio LA, et al. State of type 1 diabetes management and outcomes from the T1D Exchange in 2016-2018. *Diabetes Technol Ther* 2019 Feb;21(2):66-72 [FREE Full text] [doi: [10.1089/dia.2018.0384](https://doi.org/10.1089/dia.2018.0384)] [Medline: [30657336](https://pubmed.ncbi.nlm.nih.gov/30657336/)]
10. Gajewska KA, Bennett K, Biesma R, Sreenan S. Low uptake of continuous subcutaneous insulin infusion therapy in people with type 1 diabetes in Ireland: a retrospective cross-sectional study. *BMC Endocr Disord* 2020 Jun 23;20(1):92 [FREE Full text] [doi: [10.1186/s12902-020-00573-w](https://doi.org/10.1186/s12902-020-00573-w)] [Medline: [32576284](https://pubmed.ncbi.nlm.nih.gov/32576284/)]
11. van den Boom L, Karges B, Auzanneau M, Rami-Merhar B, Lilienthal E, von Sengbusch S, et al. Temporal trends and contemporary use of insulin pump therapy and glucose monitoring among children, adolescents, and adults with type 1 diabetes between 1995 and 2017. *Diabetes Care* 2019 Nov;42(11):2050-2056. [doi: [10.2337/dc19-0345](https://doi.org/10.2337/dc19-0345)] [Medline: [31488568](https://pubmed.ncbi.nlm.nih.gov/31488568/)]
12. DeSalvo DJ, Miller KM, Hermann JM, Maahs DM, Hofer SE, Clements MA, T1D Exchange and DPV Registries. Continuous glucose monitoring and glycemic control among youth with type 1 diabetes: International comparison from the T1D Exchange and DPV Initiative. *Pediatr Diabetes* 2018 Nov;19(7):1271-1275 [FREE Full text] [doi: [10.1111/pedi.12711](https://doi.org/10.1111/pedi.12711)] [Medline: [29923262](https://pubmed.ncbi.nlm.nih.gov/29923262/)]
13. Guldmond N, Hercheui MD. Technology care for patients with chronic conditions: The chronic care model as a framework for the integration of ICT. In: *ICT Critical Infrastructures and Society*. 2012 Presented at: IFIP International Conference on Human Choice and Computers; September 27-28, 2012; Amsterdam, The Netherlands p. 123-133. [doi: [10.1007/978-3-642-33332-3\\_12](https://doi.org/10.1007/978-3-642-33332-3_12)]
14. Vrijhoef HJ, de Belvis AG, de la Calle M, de Sabata MS, Hauck B, Montante S, et al. IT-supported integrated care pathways for diabetes: A compilation and review of good practices. *Int J Care Coord* 2017 Jun;20(1-2):26-40 [FREE Full text] [doi: [10.1177/2053434517714427](https://doi.org/10.1177/2053434517714427)] [Medline: [28690856](https://pubmed.ncbi.nlm.nih.gov/28690856/)]

15. Joshi AU, Lewiss RE, Aini M, Babula B, Henwood PC. Solving community SARS-CoV-2 testing with telehealth: Development and implementation for screening, evaluation and testing. *JMIR Mhealth Uhealth* 2020 Oct 16;8(10):e20419 [FREE Full text] [doi: [10.2196/20419](https://doi.org/10.2196/20419)] [Medline: [33006942](https://pubmed.ncbi.nlm.nih.gov/33006942/)]
16. Gujral UP, Johnson L, Nielsen J, Vellanki P, Haw JS, Davis GM, et al. Preparedness cycle to address transitions in diabetes care during the COVID-19 pandemic and future outbreaks. *BMJ Open Diabetes Res Care* 2020 Jul;8(1):e001520 [FREE Full text] [doi: [10.1136/bmjdr-2020-001520](https://doi.org/10.1136/bmjdr-2020-001520)] [Medline: [32690631](https://pubmed.ncbi.nlm.nih.gov/32690631/)]
17. Kircher SM, Mulcahy M, Kalyan A, Weldon CB, Trosman JR, Benson AB. Telemedicine in oncology and reimbursement policy during COVID-19 and beyond. *J Natl Compr Canc Netw*. Epub ahead of print 2020 Sep 30. [doi: [10.6004/jnccn.2020.7639](https://doi.org/10.6004/jnccn.2020.7639)] [Medline: [32998106](https://pubmed.ncbi.nlm.nih.gov/32998106/)]
18. Chaudhry H, Nadeem S, Mundi R. How satisfied are patients and surgeons with telemedicine in orthopaedic care during the COVID-19 Pandemic? A systematic review and meta-analysis. *Clin Orthop Relat Res* 2021 Jan 01;479(1):47-56. [doi: [10.1097/CORR.0000000000001494](https://doi.org/10.1097/CORR.0000000000001494)] [Medline: [33009231](https://pubmed.ncbi.nlm.nih.gov/33009231/)]
19. Rivas JG, Rodríguez-Serrano A, Loeb S, Teoh JYC, Ribal MJ, Bloemberg J, et al. Telemedicine and smart working: Spanish adaptation of the European Association of Urology recommendations. *Actas Urol Esp* 2020 Dec;44(10):644-652 [FREE Full text] [doi: [10.1016/j.acuro.2020.08.010](https://doi.org/10.1016/j.acuro.2020.08.010)] [Medline: [33012592](https://pubmed.ncbi.nlm.nih.gov/33012592/)]
20. Vazquez J, Islam T, Gursky J, Beller J, Correa DJ. Access to care matters: Remote health care needs during COVID-19. *Telemed J E Health*. Epub ahead of print 2020 Sep 30. [doi: [10.1089/tmj.2020.0371](https://doi.org/10.1089/tmj.2020.0371)] [Medline: [33017270](https://pubmed.ncbi.nlm.nih.gov/33017270/)]
21. Betancourt JA, Rosenberg MA, Zevallos A, Brown JR, Mileski M. The impact of COVID-19 on telemedicine utilization across multiple service lines in the United States. *Healthcare (Basel)* 2020 Oct 01;8(4):380 [FREE Full text] [doi: [10.3390/healthcare8040380](https://doi.org/10.3390/healthcare8040380)] [Medline: [33019667](https://pubmed.ncbi.nlm.nih.gov/33019667/)]
22. Mancebo G, Solé-Sedeño JM, Membrive I, Taus A, Castells M, Serrano L, et al. Gynecologic cancer surveillance in the era of SARS-CoV-2 (COVID-19). *Int J Gynecol Cancer*. Epub ahead of print 2020 Oct 05. [doi: [10.1136/ijgc-2020-001942](https://doi.org/10.1136/ijgc-2020-001942)] [Medline: [33020205](https://pubmed.ncbi.nlm.nih.gov/33020205/)]
23. Kriem J, Rahhal R. COVID-19 pandemic and challenges in pediatric gastroenterology practice. *World J Gastroenterol* 2020 Sep 28;26(36):5387-5394 [FREE Full text] [doi: [10.3748/wjg.v26.i36.5387](https://doi.org/10.3748/wjg.v26.i36.5387)] [Medline: [33024391](https://pubmed.ncbi.nlm.nih.gov/33024391/)]
24. Reinhard MA, Burkhardt G, Grosse-Wentrup F, Eser-Valerie D, Mumm FHA, Barnikol-Oettler B, et al. [Psychosocial support during the COVID-19 pandemic: interdisciplinary concept of care at a university hospital]. *Nervenarzt*. Epub ahead of print 2020 Oct 06 [FREE Full text] [doi: [10.1007/s00115-020-01014-8](https://doi.org/10.1007/s00115-020-01014-8)] [Medline: [33025071](https://pubmed.ncbi.nlm.nih.gov/33025071/)]
25. Kaushik A, Patel S, Dubey K. Digital cardiovascular care in COVID-19 pandemic: A potential alternative? *J Card Surg*. Epub ahead of print 2020 Oct 10 [FREE Full text] [doi: [10.1111/jocs.15094](https://doi.org/10.1111/jocs.15094)] [Medline: [33040399](https://pubmed.ncbi.nlm.nih.gov/33040399/)]
26. Schifeling CH, Shanbhag P, Johnson A, Atwater RC, Koljack C, Parnes BL, et al. Disparities in video and telephone visits among older adults during the COVID-19 pandemic: Cross-sectional analysis. *JMIR Aging* 2020 Nov 10;3(2):e23176 [FREE Full text] [doi: [10.2196/23176](https://doi.org/10.2196/23176)] [Medline: [33048821](https://pubmed.ncbi.nlm.nih.gov/33048821/)]
27. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health* 2020 Aug 01;20(1):1193 [FREE Full text] [doi: [10.1186/s12889-020-09301-4](https://doi.org/10.1186/s12889-020-09301-4)] [Medline: [32738884](https://pubmed.ncbi.nlm.nih.gov/32738884/)]
28. Tebeje TH, Klein J. Applications of e-Health to support person-centered health care at the time of COVID-19 Pandemic. *Telemed J E Health*. Epub ahead of print 2020 Jul 31. [doi: [10.1089/tmj.2020.0201](https://doi.org/10.1089/tmj.2020.0201)] [Medline: [32746750](https://pubmed.ncbi.nlm.nih.gov/32746750/)]
29. Roberts JP, Fisher TR, Trowbridge MJ, Bent C. A design thinking framework for healthcare management and innovation. *Healthc (Amst)* 2016 Mar;4(1):11-14. [doi: [10.1016/j.hjdsi.2015.12.002](https://doi.org/10.1016/j.hjdsi.2015.12.002)] [Medline: [27001093](https://pubmed.ncbi.nlm.nih.gov/27001093/)]
30. Shaw J, Agarwal P, Desveaux L, Palma DC, Stamenova V, Jamieson T, et al. Beyond "implementation": digital health innovation and service design. *NPJ Digit Med* 2018 Sep 28;1:48 [FREE Full text] [doi: [10.1038/s41746-018-0059-8](https://doi.org/10.1038/s41746-018-0059-8)] [Medline: [31304327](https://pubmed.ncbi.nlm.nih.gov/31304327/)]
31. Service design - Design is not just for products. Interaction Design Foundation. URL: <https://www.interaction-design.org/literature/article/service-design-design-is-not-just-for-products> [accessed 2020-09-04]
32. Alter S. Value blueprint and service design space for facilitating value creation. In: Proceedings of the Nineteenth Americas Conference on Information Systems. 2013 May 08 Presented at: The Nineteenth Americas Conference on Information Systems; August 15-17, 2013; Chicago, Illinois URL: <https://repository.usfca.edu/cgi/viewcontent.cgi?article=1018&context=at> [doi: [10.1007/springerreference\\_61782](https://doi.org/10.1007/springerreference_61782)]
33. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: Literature review and case study. *Milbank Q* 2016 Jun;94(2):392-429 [FREE Full text] [doi: [10.1111/1468-0009.12197](https://doi.org/10.1111/1468-0009.12197)] [Medline: [27265562](https://pubmed.ncbi.nlm.nih.gov/27265562/)]
34. Neinstein A, Wong J, Look H, Arbiter B, Quirk K, McCanne S, et al. A case study in open source innovation: developing the Tidepool Platform for interoperability in type 1 diabetes management. *J Am Med Inform Assoc* 2016 Mar;23(2):324-332 [FREE Full text] [doi: [10.1093/jamia/ocv104](https://doi.org/10.1093/jamia/ocv104)] [Medline: [26338218](https://pubmed.ncbi.nlm.nih.gov/26338218/)]
35. Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Text von Bedeutung für den EWR). *EUR-Lex*. URL: [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L\\_.2016.119.01.0001.01.DEU](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2016.119.01.0001.01.DEU) [accessed 2020-09-04]

36. Frielitz FS, Storm N, Hiort O, Katalinic A, von Sengbusch S. [The creation of a data protection policy: a guide to telemedicine healthcare projects]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2019 Apr;62(4):479-485. [doi: [10.1007/s00103-019-02918-w](https://doi.org/10.1007/s00103-019-02918-w)] [Medline: [30874830](#)]
37. Data protection statement. Charité – Universitätsmedizin Berlin. URL: [https://www.charite.de/en/service/data\\_protection/](https://www.charite.de/en/service/data_protection/) [accessed 2020-10-16]
38. Wurzel- und CA-Zertifikate im Sicherheitsniveau Global Generation 2. Deutsches Forschungsnetzwerk. URL: <https://www.pki.dfn.de/wurzelzertifikate/globalroot2/> [accessed 2020-10-16]
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](#)]
40. Joensen LE, Madsen KP, Holm L, Nielsen KA, Rod MH, Petersen AA, et al. Diabetes and COVID-19: psychosocial consequences of the COVID-19 pandemic in people with diabetes in Denmark-what characterizes people with high levels of COVID-19-related worries? *Diabet Med* 2020 Jul;37(7):1146-1154 [FREE Full text] [doi: [10.1111/dme.14319](https://doi.org/10.1111/dme.14319)] [Medline: [32392380](#)]
41. Dubey S, Biswas P, Ghosh R, Chatterjee S, Dubey MJ, Chatterjee S, et al. Psychosocial impact of COVID-19. *Diabetes Metab Syndr* 2020;14(5):779-788 [FREE Full text] [doi: [10.1016/j.dsx.2020.05.035](https://doi.org/10.1016/j.dsx.2020.05.035)] [Medline: [32526627](#)]
42. Guttmann-Bauman I, Kono J, Lin AL, Ramsey KL, Boston BA. Use of telehealth videoconferencing in pediatric type 1 diabetes in Oregon. *Telemed J E Health* 2018 Jan;24(1):86-88. [doi: [10.1089/tmj.2017.0072](https://doi.org/10.1089/tmj.2017.0072)] [Medline: [28654350](#)]
43. Döğler E, Bozbulut R, Acar A, Ercan Ş, Uğurlu AK, Akbaş ED, et al. Effect of telehealth system on glycemic control in children and adolescents with type 1 diabetes. *J Clin Res Pediatr Endocrinol* 2019 Feb 20;11(1):70-75 [FREE Full text] [doi: [10.4274/jcrpe.galenos.2018.2018.0017](https://doi.org/10.4274/jcrpe.galenos.2018.2018.0017)] [Medline: [30015620](#)]
44. Crossen S, Glaser N, Sauers-Ford H, Chen S, Tran V, Marcin J. Home-based video visits for pediatric patients with poorly controlled type 1 diabetes. *J Telemed Telecare* 2020 Jul;26(6):349-355. [doi: [10.1177/1357633X19828173](https://doi.org/10.1177/1357633X19828173)] [Medline: [30871408](#)]
45. Nørgaard K. Telemedicine consultations and diabetes technology during COVID-19. *J Diabetes Sci Technol* 2020 Jul;14(4):767-768 [FREE Full text] [doi: [10.1177/1932296820929378](https://doi.org/10.1177/1932296820929378)] [Medline: [32429702](#)]
46. Garg SK, Rodbard D, Hirsch IB, Forlenza GP. Managing new-onset type 1 diabetes during the COVID-19 pandemic: Challenges and opportunities. *Diabetes Technol Ther* 2020 Jun;22(6):431-439. [doi: [10.1089/dia.2020.0161](https://doi.org/10.1089/dia.2020.0161)] [Medline: [32302499](#)]
47. Fuchs J, Hovorka R. COVID-19 and diabetes: Could diabetes technology research help pave the way for remote healthcare? *J Diabetes Sci Technol* 2020 Jul;14(4):735-736 [FREE Full text] [doi: [10.1177/1932296820929714](https://doi.org/10.1177/1932296820929714)] [Medline: [32475169](#)]
48. Panziner D. Role of non-profit organizations during COVID-19 for diabetes care: Health care inequities and role of virtual specialty clinic. *Diabetes Technol Ther* 2020 Jun;22(6):440-443. [doi: [10.1089/dia.2020.0219](https://doi.org/10.1089/dia.2020.0219)] [Medline: [32392431](#)]
49. Teles M, Sacchetta T, Matsumoto Y. COVID-19 pandemic triggers telemedicine regulation and intensifies diabetes management technology adoption in Brazil. *J Diabetes Sci Technol* 2020 Jul;14(4):797-798 [FREE Full text] [doi: [10.1177/1932296820930033](https://doi.org/10.1177/1932296820930033)] [Medline: [32482098](#)]
50. Peters AL, Garg SK. The silver lining to COVID-19: Avoiding diabetic ketoacidosis admissions with telehealth. *Diabetes Technol Ther* 2020 May 05;22(6):449-453. [doi: [10.1089/dia.2020.0187](https://doi.org/10.1089/dia.2020.0187)] [Medline: [32383989](#)]
51. Jones MS, Goley AL, Alexander BE, Keller SB, Caldwell MM, Buse JB. Inpatient transition to virtual care during COVID-19 pandemic. *Diabetes Technol Ther* 2020 Jun;22(6):444-448. [doi: [10.1089/dia.2020.0206](https://doi.org/10.1089/dia.2020.0206)] [Medline: [32396395](#)]
52. Peyrot M, International DAWN Youth Survey Group. The DAWN Youth WebTalk Study: methods, findings, and implications. *Pediatr Diabetes* 2009 Dec;10 Suppl 13:37-45. [doi: [10.1111/j.1399-5448.2009.00612.x](https://doi.org/10.1111/j.1399-5448.2009.00612.x)] [Medline: [19930225](#)]
53. Wong JC, Neinstein AB, Spindler M, Adi S. A minority of patients with type 1 diabetes routinely downloads and retrospectively reviews device data. *Diabetes Technol Ther* 2015 Aug;17(8):555-562 [FREE Full text] [doi: [10.1089/dia.2014.0413](https://doi.org/10.1089/dia.2014.0413)] [Medline: [26133226](#)]
54. Arbiter B, Look H, McComb L, Snider C. Why download data: The benefits and challenges of more diabetes data. *Diabetes Spectr* 2019 Aug;32(3):221-225 [FREE Full text] [doi: [10.2337/ds18-0099](https://doi.org/10.2337/ds18-0099)] [Medline: [31462877](#)]
55. Wong JC, Izadi Z, Schroeder S, Nader M, Min J, Neinstein AB, et al. A pilot study of use of a software platform for the collection, integration, and visualization of diabetes device data by health care providers in a multidisciplinary pediatric setting. *Diabetes Technol Ther* 2018 Dec;20(12):806-816 [FREE Full text] [doi: [10.1089/dia.2018.0251](https://doi.org/10.1089/dia.2018.0251)] [Medline: [30461307](#)]
56. Tanenbaum ML, Hanes SJ, Miller KM, Naranjo D, Bensen R, Hood KK. Diabetes device use in adults with type 1 diabetes: barriers to uptake and potential intervention targets. *Diabetes Care* 2017 Feb;40(2):181-187 [FREE Full text] [doi: [10.2337/dc16-1536](https://doi.org/10.2337/dc16-1536)] [Medline: [27899489](#)]
57. Renard E. Insulin pump use in Europe. *Diabetes Technol Ther* 2010 Jun;12 Suppl 1:S29-S32. [doi: [10.1089/dia.2009.0189](https://doi.org/10.1089/dia.2009.0189)] [Medline: [20515303](#)]
58. Gajewska KA, Biesma R, Bennett K, Sreenan S. Barriers and facilitators to accessing insulin pump therapy by adults with type 1 diabetes mellitus: a qualitative study. *Acta Diabetol* 2021 Jan;58(1):93-105. [doi: [10.1007/s00592-020-01595-5](https://doi.org/10.1007/s00592-020-01595-5)] [Medline: [32862261](#)]

59. Lawton J, Kirkham J, Rankin D, White DA, Elliott J, Jaap A, REPOSE Group. Who gains clinical benefit from using insulin pump therapy? A qualitative study of the perceptions and views of health professionals involved in the Relative Effectiveness of Pumps over MDI and tured Education (REPOSE) trial *Stru. Diabet Med* 2016 Feb;33(2):243-251. [doi: [10.1111/dme.12879](https://doi.org/10.1111/dme.12879)] [Medline: [26248590](https://pubmed.ncbi.nlm.nih.gov/26248590/)]
60. Tanenbaum ML, Adams RN, Hanes SJ, Barley RC, Miller KM, Mulvaney SA, et al. Optimal use of diabetes devices: Clinician perspectives on barriers and adherence to device use. *J Diabetes Sci Technol* 2017 May;11(3):484-492 [FREE Full text] [doi: [10.1177/1932296816688010](https://doi.org/10.1177/1932296816688010)] [Medline: [28745093](https://pubmed.ncbi.nlm.nih.gov/28745093/)]
61. Crossen S, Raymond J, Neinstein A. Top 10 tips for successfully implementing a diabetes telehealth program. *Diabetes Technol Ther* 2020 Dec;22(12):920-928 [FREE Full text] [doi: [10.1089/dia.2020.0042](https://doi.org/10.1089/dia.2020.0042)] [Medline: [32191141](https://pubmed.ncbi.nlm.nih.gov/32191141/)]
62. Lind M, Polonsky W, Hirsch IB, Heise T, Bolinder J, Dahlqvist S, et al. Continuous glucose monitoring vs conventional therapy for glycemic control in adults with type 1 diabetes treated with multiple daily insulin injections: The GOLD randomized clinical trial. *JAMA* 2017 Jan 24;317(4):379-387. [doi: [10.1001/jama.2016.19976](https://doi.org/10.1001/jama.2016.19976)] [Medline: [28118454](https://pubmed.ncbi.nlm.nih.gov/28118454/)]
63. Beck RW, Riddlesworth T, Ruedy K, Ahmann A, Bergenstal R, Haller S, DIAMOND Study Group. Effect of continuous glucose monitoring on glycemic control in adults with type 1 diabetes using insulin injections: The DIAMOND randomized clinical trial. *JAMA* 2017 Jan 24;317(4):371-378. [doi: [10.1001/jama.2016.19975](https://doi.org/10.1001/jama.2016.19975)] [Medline: [28118453](https://pubmed.ncbi.nlm.nih.gov/28118453/)]
64. Šoupal J, Petruželková L, Flekač M, Pelcl T, Matoulek M, Daňková M, et al. Comparison of different treatment modalities for type 1 diabetes, including sensor-augmented insulin regimens, in 52 weeks of follow-up: A COMISAIR study. *Diabetes Technol Ther* 2016 Sep;18(9):532-538 [FREE Full text] [doi: [10.1089/dia.2016.0171](https://doi.org/10.1089/dia.2016.0171)] [Medline: [27482825](https://pubmed.ncbi.nlm.nih.gov/27482825/)]
65. Battelino T, Conget I, Olsen B, Schütz-Fuhrmann I, Hommel E, Hoogma R, SWITCH Study Group. The use and efficacy of continuous glucose monitoring in type 1 diabetes treated with insulin pump therapy: a randomised controlled trial. *Diabetologia* 2012 Dec;55(12):3155-3162 [FREE Full text] [doi: [10.1007/s00125-012-2708-9](https://doi.org/10.1007/s00125-012-2708-9)] [Medline: [22965294](https://pubmed.ncbi.nlm.nih.gov/22965294/)]
66. Bergenstal RM, Tamborlane WV, Ahmann A, Buse JB, Dailey G, Davis SN, STAR 3 Study Group. Sensor-augmented pump therapy for A1C reduction (STAR 3) study: results from the 6-month continuation phase. *Diabetes Care* 2011 Nov;34(11):2403-2405 [FREE Full text] [doi: [10.2337/dc11-1248](https://doi.org/10.2337/dc11-1248)] [Medline: [21933908](https://pubmed.ncbi.nlm.nih.gov/21933908/)]
67. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical targets for continuous glucose monitoring data interpretation: Recommendations from the international consensus on time in range. *Diabetes Care* 2019 Aug;42(8):1593-1603. [doi: [10.2337/dci19-0028](https://doi.org/10.2337/dci19-0028)] [Medline: [31177185](https://pubmed.ncbi.nlm.nih.gov/31177185/)]
68. Beck RW, Bergenstal RM, Riddlesworth TD, Kollman C, Li Z, Brown AS, et al. Validation of time in range as an outcome measure for diabetes clinical trials. *Diabetes Care* 2019 Mar;42(3):400-405 [FREE Full text] [doi: [10.2337/dc18-1444](https://doi.org/10.2337/dc18-1444)] [Medline: [30352896](https://pubmed.ncbi.nlm.nih.gov/30352896/)]
69. Beck RW, Bergenstal RM, Cheng P, Kollman C, Carlson AL, Johnson ML, et al. The relationships between time in range, hyperglycemia metrics, and HbA1c. *J Diabetes Sci Technol* 2019 Jul;13(4):614-626 [FREE Full text] [doi: [10.1177/1932296818822496](https://doi.org/10.1177/1932296818822496)] [Medline: [30636519](https://pubmed.ncbi.nlm.nih.gov/30636519/)]

## Abbreviations

- CGM:** continuous glucose monitoring
- DDC:** digital diabetes clinic
- HbA<sub>1c</sub>:** hemoglobin A<sub>1c</sub>
- HCP:** health care professional
- REDCap:** Research Electronic Data Capture

*Edited by G Eysenbach; submitted 16.09.20; peer-reviewed by S Hussain, R Lal, Z Ren, N Agarwal, C Campos-Castillo; comments to author 09.10.20; revised version received 28.10.20; accepted 08.12.20; published 05.04.21.*

### *Please cite as:*

Braune K, Boss K, Schmidt-Herzel J, Gajewska KA, Thieffry A, Schulze L, Posern B, Raile K  
*Shaping Workflows in Digital and Remote Diabetes Care During the COVID-19 Pandemic via Service Design: Prospective, Longitudinal, Open-label Feasibility Trial*  
*JMIR Mhealth Uhealth* 2021;9(4):e24374  
URL: <https://mhealth.jmir.org/2021/4/e24374>  
doi: [10.2196/24374](https://doi.org/10.2196/24374)  
PMID: [33571104](https://pubmed.ncbi.nlm.nih.gov/33571104/)

©Katarina Braune, Karina Boss, Jessica Schmidt-Herzel, Katarzyna Anna Gajewska, Axel Thieffry, Lilian Schulze, Barbara Posern, Klemens Raile. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 05.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# A Patient-Oriented App (ThessHF) to Improve Self-Care Quality in Heart Failure: From Evidence-Based Design to Pilot Study

Constantinos Bakogiannis<sup>1,2\*</sup>, MD, PhD; Anastasios Tsarouchas<sup>1,2\*</sup>, MD; Dimitrios Mouselimis<sup>1,2</sup>, MD; Charalampos Lazaridis<sup>1,2</sup>, MD; Efstratios K Theofillogianakos<sup>2</sup>, MD, PhD; Antonios Billis<sup>3</sup>, BSc, PhD; Stergios Tzikas<sup>2</sup>, MD, PhD; Nikolaos Fragakis<sup>2</sup>, MD, PhD; Panagiotis D Bamidis<sup>3</sup>, BSc, PhD; Christodoulos E Papadopoulos<sup>1,2</sup>, MD, PhD; Vassilios P Vassilikos<sup>1,2</sup>, MD, PhD

<sup>1</sup>Cardiovascular Prevention and Digital Cardiology Lab, Third Cardiology Department, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>2</sup>Third Cardiology Department, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>3</sup>Lab of Medical Physics, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

\* these authors contributed equally

**Corresponding Author:**

Constantinos Bakogiannis, MD, PhD  
Cardiovascular Prevention and Digital Cardiology Lab  
Third Cardiology Department  
Aristotle University of Thessaloniki  
Konstantinoupoleos 49  
Thessaloniki, 54642  
Greece  
Phone: 30 2310892598  
Email: [bakogianniscon@gmail.com](mailto:bakogianniscon@gmail.com)

## Abstract

**Background:** Heart failure (HF) remains a major public health challenge, while HF self-care is particularly challenging. Mobile health (mHealth)-based interventions taking advantage of smartphone technology have shown particular promise in increasing the quality of self-care among these patients, and in turn improving the outcomes of their disease.

**Objective:** The objective of this study was to co-develop with physicians, patients with HF, and their caregivers a patient-oriented mHealth app, perform usability assessment, and investigate its effect on the quality of life of patients with HF and rate of hospitalizations in a pilot study.

**Methods:** The development of an mHealth app (The Hellenic Educational Self-care and Support Heart Failure app [ThessHF app]) was evidence based, including features based on previous clinically tested mHealth interventions and selected by a panel of HF expert physicians and discussed with patients with HF. At the end of alpha development, the app was rated by mHealth experts with the Mobile Application Rating Scale (MARS). The beta version was tested by patients with HF, who rated its design and content by means of the Post-Study System Usability Questionnaire (PSSUQ). Subsequently, a prospective pilot study (THESS-HF [The Effect of a Specialized Smartphone app on Heart Failure patients' quality of self-care, quality of life and hospitalization rate]) was performed to investigate the effect of app use on patients with HF over a 3-month follow-up period. The primary endpoint was patients' quality of life, which was measured with the Kansas City Cardiomyopathy Questionnaire (KCCQ) and the 5-level EQ-5D version (EQ-5D-5L). The secondary endpoints were the European Heart Failure Self-care Behavior Scale (EHFScBS) score and the hospitalization rate.

**Results:** A systematic review of mHealth-based HF interventions and expert panel suggestions yielded 18 separate app features, most of which were incorporated into the ThessHF app. A total of 14 patients and 5 mHealth experts evaluated the app. The results demonstrated a very good user experience (overall PSSUQ score 2.37 [SD 0.63], where 1 is the best, and a median MARS score of 4.55/5). Finally, 30 patients (male: n=26, 87%) participated in the THESS-HF pilot study (mean age 68.7 [SD 12.4] years). A significant increase in the quality of self-care was noted according to the EHFScBS, which increased by 4.4% (SD 7.2%) ( $P=.002$ ). The mean quality of life increased nonsignificantly after 3 months according to both KCCQ (mean increase 5.8 [SD 15] points,  $P=.054$ ) and EQ-5D-5L (mean increase 5.6% [SD 15.6%],  $P=.06$ ) scores. The hospitalization rate for the follow-up duration was 3%.

**Conclusions:** The need for telehealth services and remote self-care management in HF is of vital importance, especially in periods such as the COVID-19 pandemic. We developed a user-friendly mHealth app to promote remote self-care support in HF. In this pilot study, the use of the ThessHF app was associated with an increase in the quality of self-care. A future multicenter study will investigate the effect of the app use on long-term outcomes in patients with HF.

(*JMIR Mhealth Uhealth* 2021;9(4):e24271) doi:[10.2196/24271](https://doi.org/10.2196/24271)

## KEYWORDS

mHealth; heart failure; smartphone app; self-care; COVID-19; patients; caregivers

## Introduction

Heart failure (HF) is a major burden on patients, negatively affecting their functional status and quality of life. The incidence of HF in the United States is estimated between 2 and 5/1000 person-years [1]. There are about 26 million patients with HF globally [2]. HF decompensations occur frequently and necessitate lengthy hospital stays [3-5], while the disease causes mortality comparable with many types of cancer [6].

Managing HF is a difficult task for clinicians, and even more so for patients, as complicated and lengthy self-care is needed. Pharmacological therapy for HF consists of several different medications with different dosing strategies, especially in patients with HF with reduced ejection fraction (HFrEF) [7]. Diuretics for avoiding fluid congestion work optimally when the dosage varies depending on clinical or imaging indicators of fluid accumulation. Experienced patients often make such adjustments themselves successfully [8,9]. When medications for frequently occurring comorbidities, such as atrial fibrillation, hypertension, and diabetes, are also taken into account, the result is a labyrinthine, ever-changing medication regimen that requires time, presence of mind, and dedication to successfully adhere to [10,11].

Lifestyle changes are also a necessary but difficult part of HF self-care. Tracking fluid intake [9], daily weighing [8], increasing physical activity [12-14], and getting vaccinations [15,16] have all been shown to improve outcomes in HF, but sustained long-term adherence to the “HF lifestyle” is almost impossible without repeated interventions by a multidisciplinary team. This usually consists of physicians, nurses, dietitians, psychologists, and exercise physiologists [17].

Mobile health (mHealth) pertains to the use of mobile communications and network technologies for health care [18]. mHealth-based implementations can be designed for use by clinicians [19], nurses, allied health professionals, caregivers, or even the patients themselves [17,20]. The medium by which mHealth is delivered to the end user used to be mobile phone technologies such as automated phone calls and SMS text messages [21], but now has mostly migrated to newer technologies, such as smartphones and tablets [22-25]. The data from different types of devices including wireless scales, blood

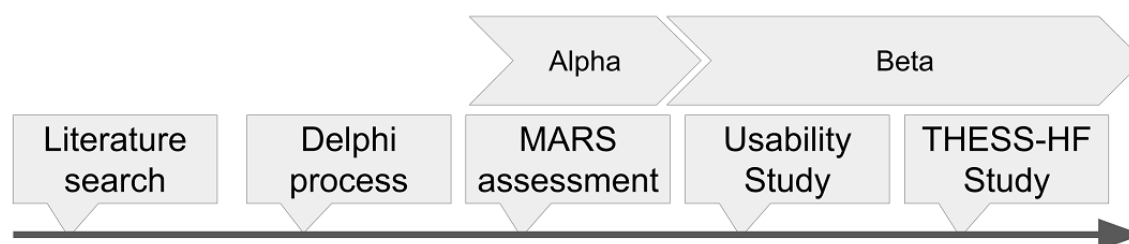
pressure monitor, or wearables (eg, sensors, bands) can be easily incorporated into such systems [26].

Patient-centered interventions utilizing mHealth technology already show promising results in improving the quality of self-care in several chronic diseases where patient participation is important, such as diabetes [27], hypertension [28], and depression [29]. The central pillar of such interventions is usually an app that provides patient education, encourages behavior that is appropriate for each disease (eg, salt restriction in HF), reminds the patient for actions that need to be taken (eg, medication/vaccination reminders), and potentially enables synchronous or asynchronous communication with health care personnel, caregivers, or even other patients with the same disease [17,22,23,25,30-32]. Remote monitoring is also a very appealing prospect [33].

In the setting of the COVID-19 pandemic, patients with severe comorbidities have a significantly higher risk of severe or even deadly disease progression [34]. It is thus of paramount importance that these patients are shielded from exposure to the virus [35]. In this situation, mHealth solutions for the remote monitoring and care of these patients may indeed become a crucial step in safeguarding this fragile group of patients. Coronavirus or not, HF, with all its intricate self-care, appears to be a prime target for mHealth-based interventions. Nonetheless, designing an app for use by patients with HF may prove challenging, as their particular needs and hindrances (eg, the mild cognitive decline HF is associated with) need to be considered.

Our aim was the development of an evidence-based, patient-oriented mHealth app (The Hellenic Educational Self-care and Support Heart Failure app [ThessHF app]) in cooperation with patients and their caregivers from our department’s Heart Failure Outpatient clinic. Furthermore, we aimed to assess ThessHF app functionality with experts as well as patients with HF in real-world settings. Finally, we conducted THESS-HF, THE Effect of a Specialized Smartphone app on Heart Failure patients’ quality of self-care, quality of life and hospitalization rate, a prospective study that investigated the clinical effect of app use on patients’ quality of life and hospitalization rate. A visual representation of this process is presented in [Figure 1](#).

**Figure 1.** A timeline of the various parts of the evidence-based app development process, along with concurrent app versions. Note that the timeframes are not represented to scale in this figure.



## Methods

### Development of a Patient-Oriented mHealth App: The ThessHF App

A systematic review of available literature was the first step in developing the app, as it was expected to reveal app features and design considerations that would increase the app's usability and benefit to patients with HF. Our search focused on mHealth interventions that utilized smartphone technology; the results were codified as discrete app features that could be incorporated into the ThessHF app. These features were disseminated to a panel of 8 cardiologists and 2 physicians with mHealth familiarity. Furthermore, these features were discussed with patients with HF and their opinion was sought.

A 3-step Delphi process was used to reach a consensus on which app features were critical, optional, or indeed unsuitable for a patient-oriented HF app. After a consensus was achieved, physicians with programming experience then developed a beta version of the app. Throughout the app's development process, patients with HF and caregivers were frequently called upon to give unstructured feedback. Prior to the app's rollout in the usability study, 5 independent mHealth experts were given access to the ThessHF app for a month and invited to rate it using the Mobile Application Rating Scale (MARS) [36].

### Usability Study of the App

Patients that visited the HF clinic were invited to install the beta version of the app to their own Android smartphones for usability testing. Patients were serially recruited from the HF clinic, with exclusion criteria being not owning a smartphone, not understanding written Greek, and denying participation. No incentive was given to participate in the usability study. After a short 30-minute hands-on session in which the researchers trained patients in the app's use, they were free to use it in any way they saw fit. A telephone number was given to them, which they could call for technical assistance or to report bugs. After a month of in-the-wild use of ThessHF, they evaluated the app by filling out the Post-Study System Usability Questionnaire (PSSUQ). The PSSUQ is a 16-item questionnaire that measures users' perceived satisfaction of a system. PSSUQ scores can range from 1 to 7, with lower scores being better [37]. Its questions can be divided in 3 subdomains: system usefulness, information quality, and interface quality.

### The Effect of a Specialized Smartphone App on Heart Failure Patients' Quality of Self-Care, Quality of Life, and Hospitalization Rate of Patients with Heart Failure (THESS-HF)

To examine the clinical effectiveness of the ThessHF app, we designed the "The Effect of a Specialized Smartphone app on Heart Failure patients' quality of self-care, quality of life and hospitalization rate" (THESS-HF) study. The THESS-HF study is a single-center, prospective study that recruits patients with HFREF who own smartphones. Because the study took place largely during the COVID-19 pandemic, it was designed to obviate the need for physical visits, with all questionnaires and patient contact in general happening via telephone, instant messaging, or video conference calls. Patients were serially recruited from our department's HF clinic. Patients should have HFREF to qualify for inclusion in the study. Exclusion criteria were cognitive or visual impairment (defined as Montreal Cognitive Assessment score <20 and visual acuity worse than 20/50, respectively), a history of stroke in the preceding 12 months, and experiencing uncontrolled psychiatric diseases. No incentive was given to participate in the study.

At baseline, HF-specific quality of life was quantified using the Kansas City Cardiomyopathy Questionnaire (KCCQ) [38], whereas health-related quality of life was quantified with the 5-level EQ-5D version (EQ-5D-5L) [39]. Quality of self-care was measured with the European Heart Failure Self-care Behavior Scale (EHFScBs) [40]. Patients then received remote instruction regarding the installation and use of the ThessHF app. Patients were then followed up on for 3 months in total. After 3 months of app use, the KCCQ, EQ-5D-5L, and EHFScBs questionnaires were once again administered. The study's primary endpoints were patients' HF-specific and health-related quality of life, as quantified via the KCCQ's total test score (KCCQ-TTS, ranging from 23 to 100, where 100 represents best HF-related quality of life) and the EQ-5D-5L visual analog scale score (EQ-5D-5L VAS, ranging from 0 to 100, where 100 represents the best health-related quality of life), respectively. Secondary endpoints were the quality of self-care score (EHFScBs, ranging from 0 to 100, where 100 represents the best quality of self-care) as well as the rate of hospitalizations or ER visits for HF decompensation during the follow-up period.

The study was approved by the local Institutional Research Board. The procedures followed were in accordance with the Helsinki Declaration of 1975, as revised in 2000.

## Statistical Analysis

Normally distributed continuous variables are reported as mean (SD) in text. Non-normally distributed continuous variables are reported as median (interquartile range). Normality was examined with the Kolmogorov–Smirnov test. The paired *t* test was used when comparing means between samples for normally distributed values, while the Mann–Whitney *U* test was used for comparing means between non-normally distributed values. The paired *t* test was used when evaluating changes in parameters for significance. Statistical significance was defined at a level  $P < .05$ . Spearman rho was used when looking for correlation between continuous or discrete variables.

## Results

### Feature Selection and App Development

Our search yielded 4 studies [22-25] that measured the effect of app-based interventions in patients with HF. The app features extracted from the literature review as presented in Table 1 were compiled and presented to the panel of experts participating in the Delphi process (a detailed review description of the systematic review is presented in Multimedia Appendix 1). The result of the process, as described in the “Methods” section, was the creation of a list of 18 critical, optional, and unsuitable/unnecessary app features (Table 2). The authors decided to declare active physician involvement a priori unsuitable, as such a feature would measurably compromise the scalability and cost-effectiveness of any mHealth intervention that used the ThessHF smartphone app.

**Table 1.** An overview of the 4 randomized controlled trials included in the systematic review.

Studies	Patients randomized to control or intervention group, study duration	App features	Hospitalizations	Quality of life	Notable outcomes	Comments
Seto et al [22] <sup>a</sup>	50/50, 6 months	<ul style="list-style-type: none"> <li>Blood pressure and weight daily (F<sup>b</sup>)</li> <li>Once a week 1-lead electrocardiography (F)</li> <li>-On-call doctor can be contacted through the app (F)</li> <li>-Active doctor involvement in setting weight targets (N<sup>c</sup>)</li> </ul>	No difference observed between groups ( $P=.1$ )	<ul style="list-style-type: none"> <li>MLHFQ<sup>d</sup> utilized</li> <li>Improved in the intervention group versus control (<math>P=.05</math>)</li> </ul>	More patients were prescribed with aldosterone antagonists in the intervention group ( $P=.02$ )	<ul style="list-style-type: none"> <li>Relatively outdated smartphone technology was utilized</li> <li>Visits to the HF<sup>e</sup> clinic and nurse workload were increased disproportionately to outcomes</li> </ul>
Vuorinen et al [23] <sup>a</sup>	47/47, 6 months	<ul style="list-style-type: none"> <li>Only buttons are used in user interface (N)</li> <li>Blood pressure, heart rate, and weight daily (F)</li> <li>Assessment of dizziness, dyspnea, palpitations, weakness, edema (F)</li> <li>Measurement history as graphs (F)</li> <li>Active nurse involvement in monitoring patient data, encouraging app use (N)</li> </ul>	No difference observed between groups ( $P=.35$ )	N/A <sup>f</sup>	<ul style="list-style-type: none"> <li>More up-titration events of angiotensin-converting enzyme inhibitor/beta blocker medication (<math>P=.04</math>) and down-titration of diuretics (<math>P=.02</math>) in the intervention group versus control</li> <li>Medical staff's (nurses) time spent for the intervention group was significantly greater versus control (<math>P&lt;.001</math>)</li> </ul>	No difference between groups regarding N-terminal pro-brain natriuretic peptide (NT-proBNP), left ventricular ejection fraction, and other clinical variables
Hägglund et al [24] <sup>g</sup>	40/32, 3 months	<ul style="list-style-type: none"> <li>Body weight via wirelessly connected scale daily (F)</li> <li>App-directed diuretics titration (F)</li> <li>Visual analog scale assessment (N)</li> <li>App-directed alert to contact HF center via phone (F)</li> <li>Patient education module (F)</li> </ul>	2.2 less hospital days per patient due to HF for the intervention group versus control (relative risk 0.38; $P<.05$ )	<ul style="list-style-type: none"> <li>KCCQ<sup>h</sup> and SF-36<sup>i</sup> utilized</li> <li>KCCQ showed significant improvement in the intervention group versus control (<math>P&lt;.05</math>)</li> </ul>	N/A	<ul style="list-style-type: none"> <li>Data were stored in the tablet</li> <li>The intervention included patient education and advices regarding self-care in adherence to the guidelines for HF (eg, consult for increase in diuretics if body weight gain detected)</li> </ul>

Studies	Patients randomized to control or intervention group, study duration	App features	Hospitalizations	Quality of life	Notable outcomes	Comments
Athilingam et al [25] <sup>a</sup>	9/9, 1 month	<ul style="list-style-type: none"> <li>Body weight daily (F)</li> <li>Symptom assessment daily (F)</li> <li>Heart rate and activity monitoring via chest strap (F)</li> <li>Medication tracker and reminder (F)</li> <li>Patient education module (F)</li> <li>Deep breathing and walking exercises (F)</li> </ul>	N/A	<ul style="list-style-type: none"> <li>KCCQ utilized</li> <li>No statistical difference</li> </ul>	Significant improvement regarding self-care management ( $P=.01$ ) and self-care confidence ( $P=.03$ ) of the intervention group versus the control group, as appraised by the Self-Care of Heart Failure Index	<ul style="list-style-type: none"> <li>Small sample size</li> <li>Only 72% of the patients concluded the 30-day follow-up</li> <li>Patients stated their preference for the use a smartphone app alone or combined with a wrist wearable tracker over the chest strap</li> </ul>

<sup>a</sup>The study intervention was smartphone based.

<sup>b</sup>Denotes functional features (ie, those directly specific to patient self-management).

<sup>c</sup>Denotes nonfunctional features (ie, those not directly specific to patient self-management features).

<sup>d</sup>MLHFQ: Minnesota Living with Heart Failure Questionnaire.

<sup>e</sup>HF: heart failure.

<sup>f</sup>N/A: not applicable.

<sup>g</sup>The study intervention was tablet based.

<sup>h</sup>KCCQ: Kansas City Cardiomyopathy Questionnaire.

<sup>i</sup>SF-36: 36-Item Short Form Survey.

**Table 2.** App features evaluated by a panel of 8 cardiologists and 2 physicians with mHealth expertise.

Critical	Optional	Unnecessary/unsuitable
<ul style="list-style-type: none"> <li>App-directed alert to contact heart failure center/medical personnel</li> <li>Blood pressure measurement</li> <li>Body weight measurement</li> <li>Buttons-only user interface</li> <li>Medication reminder</li> <li>Patient education module</li> </ul>	<ul style="list-style-type: none"> <li>Activity tracking</li> <li>Active nurse involvement in monitoring patient data</li> <li>Gamification features<sup>ab</sup></li> <li>Measurement history as graphs<sup>b</sup></li> <li>Pulses measurement<sup>b</sup></li> <li>Symptom assessment<sup>b</sup> (eg, dizziness, edem<sup>a</sup>)</li> <li>Blood glucose measurements in diabetics<sup>a, b</sup></li> </ul>	<ul style="list-style-type: none"> <li>Active doctor involvement in monitoring parameters/reacting to emergency calls</li> <li>App-directed diuretics titration</li> <li>Chest strap</li> <li>One-lead electrocardiography</li> <li>Visual analog scale assessment</li> <li>Wireless weighing scale</li> </ul>

<sup>a</sup>Features not extracted during the systematic review.

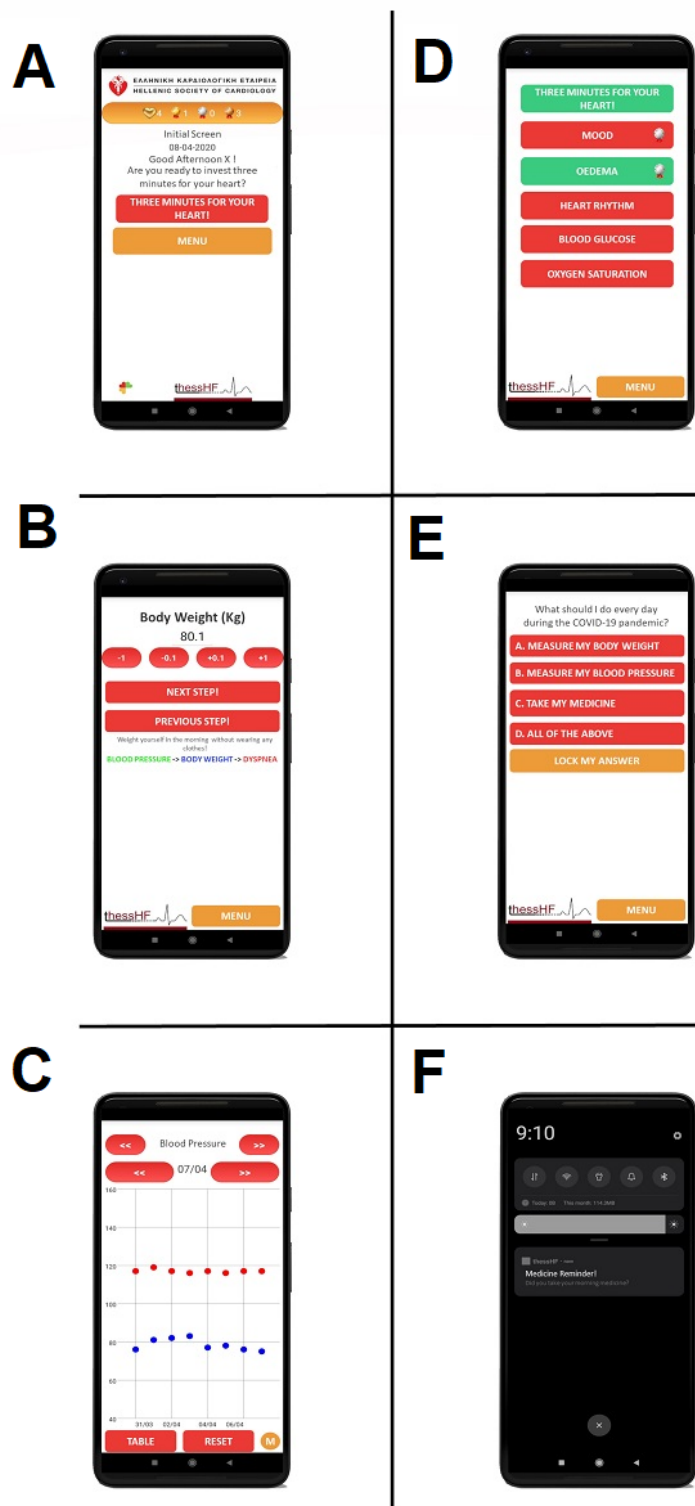
<sup>b</sup>Optional features introduced in the ThessHF app.

## The ThessHF App

Physicians developed the ThessHF app, incorporating all “critical” and most “optional” features (Table 2) in the beta version. The app encourages patients to invest 3 minutes daily to perform necessary self-care steps (weigh themselves, measure their blood pressure, and quantify potential dyspnea) and log the results in the app. Values that lie outside the ranges predetermined by clinicians alert the patient to seek medical help, while a timeline of the aforementioned parameters is always available. Patients are also reminded to take their medication in the morning, afternoon, and evening. Patient education is included in the form of a weekly quiz that contains

questions about the disease. In an effort to achieve long-term adherence to the app, gamification features have been implemented, rewarding patients with medals as they interact with them. The app features weight and symptoms tracking, medication reminders, gamification features, and a weekly HF quiz for educational purposes (Figure 2, Multimedia Appendix 2). A translated version of the text found in the app is presented in Multimedia Appendix 3. With regard to the app evaluation by the mHealth experts with the MARS, the app received a median score of 3.80 regarding user engagement, 4.0 for functionality, 4.7 for aesthetics, 4.68 for information, and 4.1 for subjective quality. The median overall score was 4.55.

**Figure 2.** Screenshots of the ThessHF app translated in English, displaying (A) the main screen and gamification features, (B) the activity where patients input their blood pressure, (C) history of blood pressure available as a graph, (D) the main menu screen, where inputted parameters turn green, (E) a question out of the weekly quiz, and (F) the push notification reminding patients to take their pills.



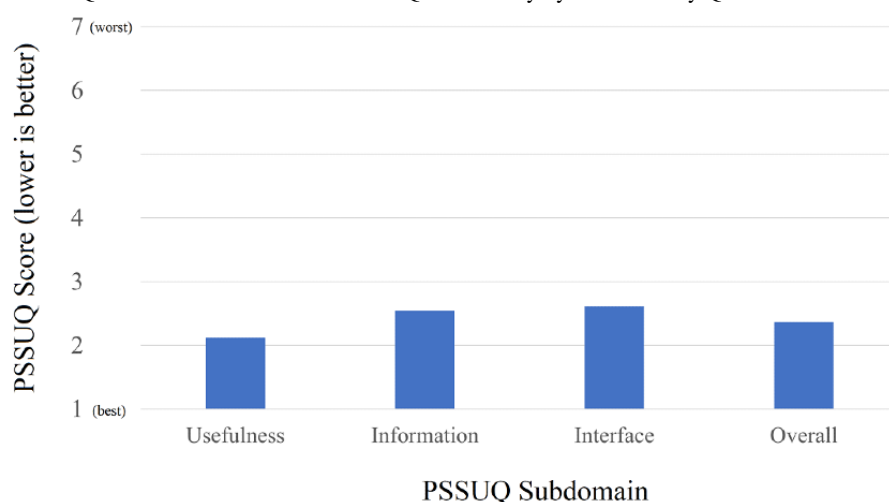
### ThessHF Usability Study

Overall, 25 patients were assessed for participating in the usability study. Among these, 11 were excluded for not owning an Android smartphone. No patient was excluded for not understanding written Greek or denying participation. In the end, 14 patients with HFrEF (mean age 64.9 [SD 9.7] years, 11

male) participated in the usability study (Table 3). The ThessHF app received an overall PSSUQ score of 2.37 (SD 0.63). In the system usefulness subdomain, the app was rated at 2.12 (SD 0.56), information quality was rated at 2.54 (SD 0.87), and the interface quality received an average score of 2.61 (SD 0.92) (Figure 3).

**Table 3.** Characteristics of patients included in the ThessHF usability.

Characteristics	Patients with heart failure with reduced ejection fraction: usability study (n=14)
Age (years), mean (SD)	64.9 (9.7)
Sex (male), n (%)	11 (79)
<b>New York Heart Association classification, n (%)</b>	
I	1 (7)
II	9 (64)
III	4 (29)
Cardiac implantable electronic devices, n (%)	8 (57)
Left ventricular ejection fraction (%), mean (SD)	27.6 (8.8)
6-Minute walking distance (m), mean (SD)	465.1 (98.3)
<b>Questionnaire answers</b>	
Smartphone ownership, n (%)	14 (100)
Confidence in smartphone use (0-5, where 5 is the best), mean (SD)	2.89 (1.81)
Social media use, n (%)	3 (21)
Primary caregiver a competent smartphone user?, n (%)	12 (86)

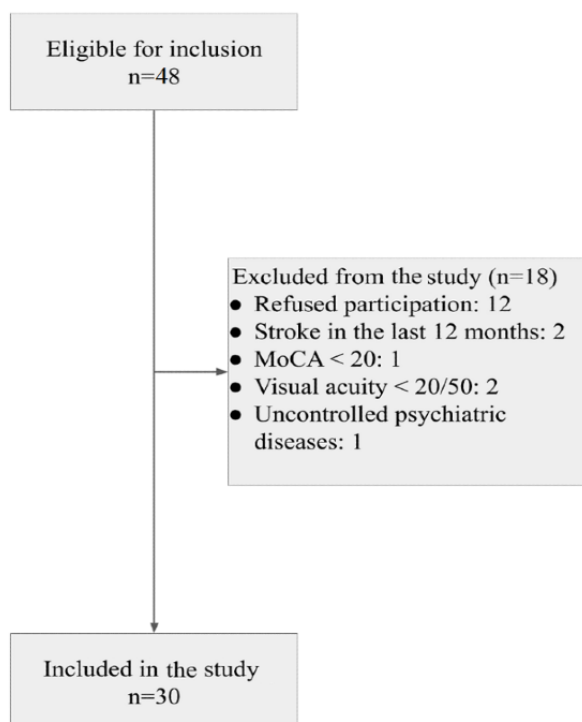
**Figure 3.** Bar chart of PSSUQ subdomain and total score. PSSUQ - Post-Study System Usability Questionnaire.

### The Results of the THESS-HF Prospective Study

A total of 30 patients were recruited in this study. The patient recruitment flowchart is presented in Figure 4. The baseline patient characteristics are tabulated in Table 4. Regarding the quality of life as measured by questionnaires KCCQ and EQ-5D-5L, a nonsignificant trend toward improvement during the study duration was observed. The mean baseline KCCQ-TTS score was 73.4 (SD 13.6), whereas the mean increase after 3

months of app usage was 5.8 (SD 15) (95% CI -0.1 to 11.6,  $P=.054$ ). The mean baseline EQ-5D-5L VAS was 59.5% (SD 14.9%), whereas the mean increase after 3 months of app usage was 5.6% (SD 15.6%) (95% CI -0.4 to 11.5,  $P=.06$ ). The mean quality of self-care significantly increased during the study duration, as the baseline EHFSBs score of 64.2% (SD 10.2%) increased by an average of 4.4% (SD 7.2%) (95% CI 1.7-7.1,  $P=.002$ ). Overall, only 1 patient was hospitalized for HF decompensation during the follow-up period.



**Figure 4.** THESS-HF study patient recruitment flowchart. MoCA – Montreal Cognitive Assessment test.**Table 4.** Baseline characteristics of patients recruited in the THESS-HF<sup>a</sup> study (n=30).

Patient parameter	Value
Age, mean (SD)	68.7 (12.4)
Male sex, n (%)	26 (87)
<b>Heart failure etiology, n (%)</b>	
Dilatated cardiomyopathy	15 (50)
Ischemic cardiomyopathy	15 (50)
Hypertension	10 (33)
Diabetes mellitus type II	12 (40)
Coronary artery disease	17 (57)
Atrial fibrillation	16 (53)
<b>Heart failure therapy at baseline, n (%)</b>	
Angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker	25 (83)
Angiotensin receptor neprilysin inhibitor	15 (50)
Beta blocker	28 (93)
Aldosterone inhibitor	24 (80)
Loop diuretics	21 (70)
Implanted cardiac implantable electronic device	14 (47)

<sup>a</sup>THESS-HF: THE Effect of a Specialized Smartphone app on Heart Failure patients' quality of life, quality of self-care and hospitalization rate.

## Discussion

### Principal Findings

In this paper we describe the entire process of designing a smartphone-based HF intervention in an evidence-based manner. Based on our team's thorough literature research, 4 trials [22-25] assessing apps' effects on patient outcomes were found. Indeed,

it seems that only a limited number of HF-specific apps have been tested through scientific studies [41]. All studies employed a similar type of intervention, asking patients to weigh themselves, measure their blood pressure, and assess their symptoms regularly, logging the results in an app. Patients found most features subjectively helpful, but it is unknown whether

some modules are more important than others in the effort to improve patient education and self-care.

Seto et al [22] were the first to develop a smartphone app for patients with HF. Monitoring patient symptom severity, blood pressure, and body weight were basic app features. These features became a staple of patient-centered HF apps, because of their significance in the disease. A substantial increase in physician and nurse workload was a recurrent motif found in many of the examined studies [23-25]. To preserve scalability of ThessHF-based interventions, we opted against any sort of in-app communication between patients and medical personnel. As an evidence-based method of intervention, the “information–motivation–behavioral skills” model, conceptualized by Fisher et al [42] and employed by Athilingam et al [25] in their HeartMapp app, served as a robust framework that allowed us to combine the features that were included in ThessHF. Our app’s advantage over previous implementations mostly lied in its evidence-based development, as well as the innovative gamification features such as the HF knowledge quiz and participation trophies, which are expected to motivate patients maintain frequent interaction with the app.

We believe that ThessHF greatly benefited from this systematic search for app features in the literature, as well as the systematized expert input in deciding which app features should be included. Patients and mHealth experts alike rated the app positively during usability testing. Some features classified as “inclusion optional” by experts were not included. Activity tracking was decided against because of the low market penetration of wearables among the population of patients with HF. The active involvement of nursing personnel in the app was not implemented to maintain scalability in resource-scarce settings. According to the PSSUQ scores, patients found the app actually useful for their everyday self-care, while experts filling out the MARS score lauded the app’s stability and quality of information. In fact, almost all PSSUQ scores, except for the interface quality subdomain, were better than the mean ratings collected by Lewis et al [43] in a psychometric study of the questionnaire. That said, both groups found that the interface could benefit from some improvements. Thanks to input from our patients, we made several design changes to suit the app to their needs. In particular, buttons were made significantly larger and spaced out, and patients were no longer required to use the onscreen keyboard for daily use. Indeed, it is likely that designing elegant yet easy-to-use apps for patients with HF will prove a major challenge for similar mHealth interventions worldwide. As mentioned above, unstructured comments made by patients and physicians trying out ThessHF were crucial to the improvement of the graphical interface. The fact that the very physicians treating these patients with HF were tasked with actually coding the desired design changes reinforced the feeling of patients that they were an active part of the development process, increasing their motivation and bonding with the HF clinic.

The effect of app use on quality of self-care and quality of life of patients with HF was investigated in the THESS-HF pilot study. Although neither HF-specific nor health-related quality of life increased significantly during the follow-up period, the quality of self-care significantly improved. Self-care quality

has been repeatedly [44-46] found to correlate with outcomes in HF. Thus, it is not unreasonable to assume that improving self-care can improve HF quality of life and morbidity in the long run. This will be tested in a multicenter prospective study aiming to recruit more patients with a longer follow-up, so as to test whether sustained use of a patient-centered HF app can yield better outcomes.

mHealth interventions are expected to become highly valuable tools in the remote care of patients with HF during the COVID-19 pandemic that started in 2019 [47]. It is thus fortuitous that innovative and noninvasive means of monitoring patients with HF have been proposed [48]. The ThessHF app proved to be a competent tool for patient support in times of this crisis. The lack of systematized patient–physician interaction or transmission of locally saved patient data through the app constituted a conscious choice to enable low physician workload and thus affordable scalability. This design decision was also advantageous insofar that patient data were never uploaded to remote databases, which would pose a significant medicolegal challenge. By contrast, physicians having direct access to patient telemetry could have further increased the app’s efficacy in improving the standard of care of patients with HF. Furthermore, usage statistics that could help highlight which features saw most use by patients as well as overall interaction time were unavailable. Our research team is actively looking into including this functionality in the future versions of the app.

As stated above, the inherent inability of data transmission through the app precluded telemetry, which could have been used to better understand different users’ engagement pattern, as well as potential technical issues that prevented patients from making full use of the app. Nonetheless, subsequent physical visits allowed us to peruse participants’ app history. This unstructured “final visit” allowed the research team to access usage statistics for some patients. Anecdotally, the research team observed that most patients replaced their handwritten arterial pressure and body weight journals with the app, a fact that kept patient interaction with the app high throughout and beyond the study duration. It should also be noted that no patient dropped out of the study, with all of them continuing some use of the app.

In many parts of the world where social distancing is sternly encouraged [35], primary care physicians and outpatient HF clinics are predicted to cease or reduce noncritical visits, whether willingly or at the behest of the local health authorities. The time originally allocated to physical visits could instead be invested in managing patients remotely, via tailor-made platforms. Indeed, now may be the time to embrace platforms with increased physician involvement, even though such attempts yielded mixed results in the past [23-25]. In this spirit, the ThessHF app is planned to incorporate remote, secure data transfer, ideally paired with wireless sensors.

In parallel with the THESS-HF study and during the lockdown imposed in Greece between March and May 2020, the Hellenic Society of Cardiology made the app available for download for all Greek patients with HF via its website, to assist them with self-care during the COVID-19 pandemic, and beyond. As of the writing of this article, 405 patients with HF downloaded the

app. Because of our longstanding commitment to respect patients' sensitive data, no analytics were available regarding app usage. Thus, the user was not asked for any data prior to downloading the app through the internet, as this would be an entirely different study.

### Conclusion

HF is a chronic disease, in which consistent and complex self-care is required to achieve good outcomes. mHealth-based interventions to improve patient education and the quality of self-care appear promising, at least in part due to the ease and

low cost of their implementation [1,3,4,49]. The first step in the evidence-based development of ThessHF was a systematic review and appraisal of similar interventions. Beneficial app features were selected by a panel of physicians and implemented by physician-programmers in the beta version of the ThessHF app, which received positive reviews by patients and mHealth experts alike. In the THESS-HF study, app use correlated with improved self-care on the part of patients. A future multicenter study with longer follow-up duration will investigate whether improvements in self-care achieved through app use can lead to better outcomes for patients with HF.

### Acknowledgments

We thank the experts that evaluated the ThessHF app with the MARS test, as well as the cardiologists who participated in the Delphi process. Furthermore, we want to thank all the patients and their caregivers that actively participated in the development and improvement of the ThessHF app. We finally want to thank the Hellenic Society of Cardiology for making the app widely available through the Society's website during the COVID-19 lockdown.

### Authors' Contributions

CB, AT, DM, ET, and PB performed the literature research, the app development, the statistical analysis, and the preparation of the draft manuscript. CB, CL, VV, ST, CP, NF, and AB contributed to design the app development process, the final study design, and the study execution. CB, AT, DM, ET, PB, and VV did the proofreading of the manuscript and suggested critical changes. All authors approve the final version of the review.

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

Detailed description of the methods, PRISMA flow diagram and Cochrane's Risk of Bias table for the systematic review that was performed to find HF-oriented mHealth apps.

[DOCX File, 68 KB - [mhealth\\_v9i4e24271\\_app1.docx](#) ]

#### Multimedia Appendix 2

Screenshots of the original ThessHF app in Greek. The figure displays (a) the main screen and gamification features, (b) the activity where patients input their blood pressure, (c) history of blood pressure available as a graph, (d) the main menu screen, where inputted parameters turn green, (e) a question out of the weekly quiz, and (f) the push notification reminding patients to take their pills.

[DOCX File, 340 KB - [mhealth\\_v9i4e24271\\_app2.docx](#) ]

#### Multimedia Appendix 3

The ThessHF app in action. The user inputs their arterial pressure, body weight and dyspnea, looks at past measurements, takes a quick quiz on HF and finally cashes in their trophies to get golden hearts.

[MP4 File (MP4 Video), 15559 KB - [mhealth\\_v9i4e24271\\_app3.mp4](#) ]

### References

1. DeFrances CJ, Hall MJ. 2005 National Hospital Discharge Survey. *Adv Data* 2007 Jul 12(385):1-19. [Medline: [17691217](#)]
2. Ponikowski P, Anker SD, AlHabib KF, Cowie MR, Force TL, Hu S, et al. Heart failure: preventing disease and death worldwide. *ESC Heart Fail* 2014 Sep;1(1):4-25. [doi: [10.1002/ehf2.12005](#)] [Medline: [28834669](#)]
3. Savarese G, Lund LH. Global Public Health Burden of Heart Failure. *Card Fail Rev* 2017 Apr;3(1):7-11. [doi: [10.15420/cfr.2016.25.2](#)] [Medline: [28785469](#)]
4. Bloom MW, Greenberg B, Jaarsma T, Januzzi JL, Lam CSP, Maggioni AP, et al. Heart failure with reduced ejection fraction. *Nat Rev Dis Primers* 2017 Aug 24;3(1):17058. [doi: [10.1038/nrdp.2017.58](#)] [Medline: [28836616](#)]
5. Bakogiannis C, Briasoulis A, Mouselimis D, Tsarouchas A, Papageorgiou N, Papadopoulos C, et al. Iron deficiency as therapeutic target in heart failure: a translational approach. *Heart Fail Rev* 2020 Mar;25(2):173-182. [doi: [10.1007/s10741-019-09815-z](#)] [Medline: [31230175](#)]

6. Stewart S, Ekman I, Ekman T, Odén A, Rosengren A. Population Impact of Heart Failure and the Most Common Forms of Cancer. *Circ Cardiovasc Qual Outcomes* 2010 Nov;3(6):573-580. [doi: [10.1161/circoutcomes.110.957571](https://doi.org/10.1161/circoutcomes.110.957571)]
7. McMurray JJV, Adamopoulos S, Anker S, Auricchio A, Böhm M, Dickstein K, ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012 Jul;33(14):1787-1847. [doi: [10.1093/eurheartj/ehs104](https://doi.org/10.1093/eurheartj/ehs104)] [Medline: [22611136](https://pubmed.ncbi.nlm.nih.gov/22611136/)]
8. Jones CD, Holmes GM, Dewalt DA, Erman B, Broucksou K, Hawk V, et al. Is adherence to weight monitoring or weight-based diuretic self-adjustment associated with fewer heart failure-related emergency department visits or hospitalizations? *J Card Fail* 2012 Jul;18(7):576-584 [FREE Full text] [doi: [10.1016/j.cardfail.2012.05.004](https://doi.org/10.1016/j.cardfail.2012.05.004)] [Medline: [22748492](https://pubmed.ncbi.nlm.nih.gov/22748492/)]
9. Johansson P, van der Wal MH, Strömberg A, Waldréus N, Jaarsma T. Fluid restriction in patients with heart failure: how should we think? *Eur J Cardiovasc Nurs* 2016 Aug 11;15(5):301-304. [doi: [10.1177/1474515116650346](https://doi.org/10.1177/1474515116650346)] [Medline: [27169459](https://pubmed.ncbi.nlm.nih.gov/27169459/)]
10. Triposkiadis FK, Skoularigis J. Prevalence and importance of comorbidities in patients with heart failure. *Curr Heart Fail Rep* 2012 Dec 24;9(4):354-362. [doi: [10.1007/s11897-012-0110-z](https://doi.org/10.1007/s11897-012-0110-z)] [Medline: [22918666](https://pubmed.ncbi.nlm.nih.gov/22918666/)]
11. Polytarchou K, Dimitroglou Y, Varvarousis D, Christodoulis N, Psachoulia C, Pantziou C, et al. Methylmalonic acid and vitamin B12 in patients with heart failure. *Hellenic J Cardiol* 2020 Sep;61(5):330-337 [FREE Full text] [doi: [10.1016/j.hjc.2019.10.010](https://doi.org/10.1016/j.hjc.2019.10.010)] [Medline: [31740360](https://pubmed.ncbi.nlm.nih.gov/31740360/)]
12. Lee H, Boo S, Yu J, Suh SR, Chun KJ, Kim J. Physical Functioning, Physical Activity, Exercise Self-Efficacy, and Quality of Life Among Individuals With Chronic Heart Failure in Korea: A Cross-Sectional Descriptive Study. *J Nurs Res* 2017 Apr;25(2):131-139. [doi: [10.1097/JNR.000000000000150](https://doi.org/10.1097/JNR.000000000000150)] [Medline: [28277393](https://pubmed.ncbi.nlm.nih.gov/28277393/)]
13. Waring T, Gross K, Soucier R, ZuWallack R. Measured Physical Activity and 30-Day Rehospitalization in Heart Failure Patients. *J Cardiopulm Rehabil Prev* 2017 Mar;37(2):124-129. [doi: [10.1097/HCR.000000000000204](https://doi.org/10.1097/HCR.000000000000204)] [Medline: [27798506](https://pubmed.ncbi.nlm.nih.gov/27798506/)]
14. Bonios MJ, Adamopoulos SN, Drakos SG. The evolving field of nonpharmacological therapies to improve functional capacity in chronic heart failure. *Hellenic J Cardiol* 2018 Nov;59(6):321-322 [FREE Full text] [doi: [10.1016/j.hjc.2018.11.001](https://doi.org/10.1016/j.hjc.2018.11.001)] [Medline: [30448617](https://pubmed.ncbi.nlm.nih.gov/30448617/)]
15. Kadoglou NPE, Bracke F, Simmers T, Tsiodras S, Parissis J. Influenza infection and heart failure-vaccination may change heart failure prognosis? *Heart Fail Rev* 2017 May 15;22(3):329-336 [FREE Full text] [doi: [10.1007/s10741-017-9614-7](https://doi.org/10.1007/s10741-017-9614-7)] [Medline: [28508293](https://pubmed.ncbi.nlm.nih.gov/28508293/)]
16. Vardeny O, Claggett B, Udell JA, Packer M, Zile M, Rouleau J, PARADIGM-HF Investigators. Influenza Vaccination in Patients With Chronic Heart Failure: The PARADIGM-HF Trial. *JACC Heart Fail* 2016 Feb;4(2):152-158 [FREE Full text] [doi: [10.1016/j.jchf.2015.10.012](https://doi.org/10.1016/j.jchf.2015.10.012)] [Medline: [26746371](https://pubmed.ncbi.nlm.nih.gov/26746371/)]
17. Athilingam P, Jenkins B. Mobile Phone Apps to Support Heart Failure Self-Care Management: Integrative Review. *JMIR Cardio* 2018 May 02;2(1):e10057 [FREE Full text] [doi: [10.2196/10057](https://doi.org/10.2196/10057)] [Medline: [31758762](https://pubmed.ncbi.nlm.nih.gov/31758762/)]
18. Istepanian R, Laxminarayan S, Pattichis C. M-Health. Istepanian RSH, Laxminarayan S, Pattichis CS. editors. Boston, MA: Springer US; 2006.
19. Atkinson SA, Bryan MH, Anderson GH. Human milk feeding in premature infants: Protein, fat, and carbohydrate balances in the first two weeks of life. *The Journal of Pediatrics* 1981 Oct;99(4):617-624. [doi: [10.1016/s0022-3476\(81\)80275-2](https://doi.org/10.1016/s0022-3476(81)80275-2)] [Medline: [7277107](https://pubmed.ncbi.nlm.nih.gov/7277107/)]
20. Tripoliti EE, Karanasiou GS, Kalatzis FG, Naka KK, Fotiadis DI. The Evolution of mHealth Solutions for Heart Failure Management. *Adv Exp Med Biol* 2018;1067:353-371. [doi: [10.1007/5584\\_2017\\_99](https://doi.org/10.1007/5584_2017_99)] [Medline: [28980271](https://pubmed.ncbi.nlm.nih.gov/28980271/)]
21. Yudi MB, Clark DJ, Tsang D, Jelinek M, Kalten K, Joshi S, et al. SMARTphone-based, early cardiac REHAbilitation in patients with acute coronary syndromes [SMART-REHAB Trial]: a randomized controlled trial protocol. *BMC Cardiovasc Disord* 2016 Sep 05;16(1):170 [FREE Full text] [doi: [10.1186/s12872-016-0356-6](https://doi.org/10.1186/s12872-016-0356-6)] [Medline: [27596569](https://pubmed.ncbi.nlm.nih.gov/27596569/)]
22. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. *J Med Internet Res* 2012 Feb 16;14(1):e31 [FREE Full text] [doi: [10.2196/jmir.1909](https://doi.org/10.2196/jmir.1909)] [Medline: [22356799](https://pubmed.ncbi.nlm.nih.gov/22356799/)]
23. Vuorinen A, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van GM, et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. *J Med Internet Res* 2014;16(12):e282 [FREE Full text] [doi: [10.2196/jmir.3651](https://doi.org/10.2196/jmir.3651)] [Medline: [25498992](https://pubmed.ncbi.nlm.nih.gov/25498992/)]
24. Hägglund E, Lyngå P, Frie F, Ullman B, Persson H, Melin M, et al. Patient-centred home-based management of heart failure. Findings from a randomised clinical trial evaluating a tablet computer for self-care, quality of life and effects on knowledge. *Scand Cardiovasc J* 2015 Aug;49(4):193-199. [doi: [10.3109/14017431.2015.1035319](https://doi.org/10.3109/14017431.2015.1035319)] [Medline: [25968968](https://pubmed.ncbi.nlm.nih.gov/25968968/)]
25. Athilingam P, Jenkins B, Johansson M, Labrador M. A Mobile Health Intervention to Improve Self-Care in Patients With Heart Failure: Pilot Randomized Control Trial. *JMIR Cardio* 2017 Aug 11;1(2):e3 [FREE Full text] [doi: [10.2196/cardio.7848](https://doi.org/10.2196/cardio.7848)] [Medline: [31758759](https://pubmed.ncbi.nlm.nih.gov/31758759/)]

26. Brooks GC, Vittinghoff E, Iyer S, Tandon D, Kuhar P, Madsen KA, et al. Accuracy and Usability of a Self-Administered 6-Minute Walk Test Smartphone Application. *Circ Heart Fail* 2015 Sep;8(5):905-913 [FREE Full text] [doi: [10.1161/CIRCHEARTFAILURE.115.002062](https://doi.org/10.1161/CIRCHEARTFAILURE.115.002062)] [Medline: [26283292](https://pubmed.ncbi.nlm.nih.gov/26283292/)]
27. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and Beliefs About Mobile Phone Apps for Diabetes Self-Management: Surveys of People in a Hospital Diabetes Clinic and Diabetes Health Professionals in New Zealand. *JMIR Mhealth Uhealth* 2017 Jun 30;5(6):e85 [FREE Full text] [doi: [10.2196/mhealth.7263](https://doi.org/10.2196/mhealth.7263)] [Medline: [28666975](https://pubmed.ncbi.nlm.nih.gov/28666975/)]
28. Kang H, Park H. A Mobile App for Hypertension Management Based on Clinical Practice Guidelines: Development and Deployment. *JMIR Mhealth Uhealth* 2016 Feb;4(1):e12 [FREE Full text] [doi: [10.2196/mhealth.4966](https://doi.org/10.2196/mhealth.4966)] [Medline: [26839283](https://pubmed.ncbi.nlm.nih.gov/26839283/)]
29. 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](https://doi.org/10.2196/jmir.6482)] [Medline: [27998876](https://pubmed.ncbi.nlm.nih.gov/27998876/)]
30. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, MOBITEL Investigators. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. *J Med Internet Res* 2009 Aug 17;11(3):e34 [FREE Full text] [doi: [10.2196/jmir.1252](https://doi.org/10.2196/jmir.1252)] [Medline: [19687005](https://pubmed.ncbi.nlm.nih.gov/19687005/)]
31. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, SUPPORT-HF Investigators. A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study. *Int J Med Inform* 2015 Oct;84(10):743-753. [doi: [10.1016/j.ijmedinf.2015.05.003](https://doi.org/10.1016/j.ijmedinf.2015.05.003)] [Medline: [26037921](https://pubmed.ncbi.nlm.nih.gov/26037921/)]
32. Kotooka N, Kitakaze M, Nagashima K, Asaka M, Kinugasa Y, Nochioka K, HOMES-HF study investigators. The first multicenter, randomized, controlled trial of home telemonitoring for Japanese patients with heart failure: home telemonitoring study for patients with heart failure (HOMES-HF). *Heart Vessels* 2018 Aug;33(8):866-876. [doi: [10.1007/s00380-018-1133-5](https://doi.org/10.1007/s00380-018-1133-5)] [Medline: [29450689](https://pubmed.ncbi.nlm.nih.gov/29450689/)]
33. Abraham WT, Adamson PB, Bourge RC, Aaron MF, Costanzo MR, Stevenson LW, et al. Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. *Lancet* 2011 Feb 19;377(9766):658-666. [doi: [10.1016/S0140-6736\(11\)60101-3](https://doi.org/10.1016/S0140-6736(11)60101-3)] [Medline: [21315441](https://pubmed.ncbi.nlm.nih.gov/21315441/)]
34. Liu W, Tao Z, Wang L, Yuan M, Liu K, Zhou L, et al. Analysis of factors associated with disease outcomes in hospitalized patients with 2019 novel coronavirus disease. *Chin Med J (Engl)* 2020 May 05;133(9):1032-1038 [FREE Full text] [doi: [10.1097/CM9.0000000000000775](https://doi.org/10.1097/CM9.0000000000000775)] [Medline: [32118640](https://pubmed.ncbi.nlm.nih.gov/32118640/)]
35. Stein R. COVID-19 and rationally layered social distancing. *Int J Clin Pract* 2020 Jul 14;74(7):e13501 [FREE Full text] [doi: [10.1111/ijcp.13501](https://doi.org/10.1111/ijcp.13501)] [Medline: [32170898](https://pubmed.ncbi.nlm.nih.gov/32170898/)]
36. 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;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
37. Lewis JR. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors Society Annual Meeting* 1992;36(12):1259-1260. [doi: [10.1177/154193129203601617](https://doi.org/10.1177/154193129203601617)]
38. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000 Apr;35(5):1245-1255 [FREE Full text] [Medline: [10758967](https://pubmed.ncbi.nlm.nih.gov/10758967/)]
39. Kontodimopoulos N, Pappa E, Niakas D, Yfantopoulos J, Dimitrakaki C, Tountas Y. Validity of the EuroQoL (EQ-5D) instrument in a Greek general population. *Value Health* 2008 Dec;11(7):1162-1169 [FREE Full text] [doi: [10.1111/j.1524-4733.2008.00356.x](https://doi.org/10.1111/j.1524-4733.2008.00356.x)] [Medline: [18489492](https://pubmed.ncbi.nlm.nih.gov/18489492/)]
40. Lambrinou E, Kalogirou F, Lamnisis D, Papathanassoglou E, Protopapas A, Sourtzi P, et al. The Greek version of the 9-item European Heart Failure Self-care Behaviour Scale: a multidimensional or a uni-dimensional scale? *Heart Lung* 2014;43(6):494-499. [doi: [10.1016/j.hrtlng.2014.07.001](https://doi.org/10.1016/j.hrtlng.2014.07.001)] [Medline: [25109661](https://pubmed.ncbi.nlm.nih.gov/25109661/)]
41. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS). *JMIR Mhealth Uhealth* 2016 Jun 14;4(2):e74 [FREE Full text] [doi: [10.2196/mhealth.5882](https://doi.org/10.2196/mhealth.5882)] [Medline: [27302310](https://pubmed.ncbi.nlm.nih.gov/27302310/)]
42. Fisher W, Fisher J, Harman J. The Information-Motivation-Behavioural Skills Model: A General Social Psychological Approach to Understanding and Promoting Health Behaviour in *Social Psychological Foundations of Health and Illness*. In: *Social Psychological Foundations of Health and Illness*. New Jersey, US: Blackwell Publishing Ltd; Jan 01, 2003:82-106.
43. Lewis JR. Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction* 2002 Sep;14(3-4):463-488. [doi: [10.1080/10447318.2002.9669130](https://doi.org/10.1080/10447318.2002.9669130)]
44. Fitzgerald AA, Powers JD, Ho PM, Maddox TM, Peterson PN, Allen LA, et al. Impact of medication nonadherence on hospitalizations and mortality in heart failure. *J Card Fail* 2011 Aug;17(8):664-669. [doi: [10.1016/j.cardfail.2011.04.011](https://doi.org/10.1016/j.cardfail.2011.04.011)] [Medline: [21807328](https://pubmed.ncbi.nlm.nih.gov/21807328/)]
45. Buck HG, Lee CS, Moser DK, Albert NM, Lennie T, Bentley B, et al. Relationship between self-care and health-related quality of life in older adults with moderate to advanced heart failure. *J Cardiovasc Nurs* 2012;27(1):8-15. [doi: [10.1097/JCN.0b013e3182106299](https://doi.org/10.1097/JCN.0b013e3182106299)] [Medline: [21558868](https://pubmed.ncbi.nlm.nih.gov/21558868/)]

46. Wu J, Frazier SK, Rayens MK, Lennie TA, Chung ML, Moser DK. Medication adherence, social support, and event-free survival in patients with heart failure. *Health Psychol* 2013 Jun;32(6):637-646 [FREE Full text] [doi: [10.1037/a0028527](https://doi.org/10.1037/a0028527)] [Medline: [22746258](https://pubmed.ncbi.nlm.nih.gov/22746258/)]
47. Velavan TP, Meyer CG. The COVID-19 epidemic. *Trop Med Int Health* 2020 Mar;25(3):278-280 [FREE Full text] [doi: [10.1111/tmi.13383](https://doi.org/10.1111/tmi.13383)] [Medline: [32052514](https://pubmed.ncbi.nlm.nih.gov/32052514/)]
48. Conn NJ, Schwarz KQ, Borkholder DA. In-Home Cardiovascular Monitoring System for Heart Failure: Comparative Study. *JMIR Mhealth Uhealth* 2019 Jan 18;7(1):e12419 [FREE Full text] [doi: [10.2196/12419](https://doi.org/10.2196/12419)] [Medline: [30664492](https://pubmed.ncbi.nlm.nih.gov/30664492/)]
49. Mosterd A, Hoes AW. Clinical epidemiology of heart failure. *Heart* 2007 Sep;93(9):1137-1146 [FREE Full text] [doi: [10.1136/hrt.2003.025270](https://doi.org/10.1136/hrt.2003.025270)] [Medline: [17699180](https://pubmed.ncbi.nlm.nih.gov/17699180/)]

## Abbreviations

**EHFScBs:** European Heart Failure Self-care Behavior Scale

**HF:** heart failure

**HFrEF:** HF with reduced ejection fraction

**KCCQ:** Kansas City Cardiomyopathy Questionnaire

**PSSUQ:** Post-Study System Usability Questionnaire

**ThessHF app:** The Hellenic Educational Self-care and Support Heart Failure app

**THESS-HF:** THE Effect of a Specialized Smartphone app on Heart Failure patients' quality of self-care, quality of life and hospitalization rate

*Edited by L Buis; submitted 11.09.20; peer-reviewed by D Zisis, A Triantafyllidis; comments to author 31.10.20; revised version received 05.12.20; accepted 02.03.21; published 13.04.21.*

*Please cite as:*

*Bakogiannis C, Tsarouchas A, Mouselimis D, Lazaridis C, Theofillogianakos EK, Billis A, Tzikas S, Fragakis N, Bamidis PD, Papadopoulos CE, Vassilikos VP*

*A Patient-Oriented App (ThessHF) to Improve Self-Care Quality in Heart Failure: From Evidence-Based Design to Pilot Study*  
*JMIR Mhealth Uhealth* 2021;9(4):e24271

URL: <https://mhealth.jmir.org/2021/4/e24271>

doi: [10.2196/24271](https://doi.org/10.2196/24271)

PMID: [33847599](https://pubmed.ncbi.nlm.nih.gov/33847599/)

©Constantinos Bakogiannis, Anastasios Tsarouchas, Dimitrios Mouselimis, Charalampos Lazaridis, Efstratios K Theofillogianakos, Antonios Billis, Stergios Tzikas, Nikolaos Fragakis, Panagiotis D Bamidis, Christodoulos E Papadopoulos, Vassilios P Vassilikos. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 13.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Usability of a Mobile App for Real-Time Assessment of Fatigue and Related Symptoms in Patients With Multiple Sclerosis: Observational Study

Miklos Palotai<sup>1</sup>, MD; Max Wallack<sup>1</sup>, MD; Gergo Kujbus<sup>2</sup>, MSc; Adam Dalnoki<sup>2</sup>, MSc; Charles Guttmann<sup>1</sup>, MD

<sup>1</sup>Center for Neurological Imaging, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

<sup>2</sup>Mobilengine, Budapest, Hungary

**Corresponding Author:**

Miklos Palotai, MD

Center for Neurological Imaging, Department of Radiology

Brigham and Women's Hospital, Harvard Medical School

1249 Boylston Street

Boston, MA, 02215

United States

Phone: 1 617 278 0613

Email: [palotai@bwh.harvard.edu](mailto:palotai@bwh.harvard.edu)

## Abstract

**Background:** Although fatigue is one of the most debilitating symptoms in patients with multiple sclerosis (MS), its pathogenesis is not well understood. Neurogenic, inflammatory, endocrine, and metabolic mechanisms have been proposed. Taking into account the temporal dynamics and comorbid mood symptoms of fatigue may help differentiate fatigue phenotypes. These phenotypes may reflect different pathogeneses and may respond to different mechanism-specific treatments. Although several tools have been developed to assess various symptoms (including fatigue), monitor clinical status, or improve the perceived level of fatigue in patients with MS, options for a detailed, real-time assessment of MS-related fatigue and relevant comorbidities are still limited.

**Objective:** This study aims to present a novel mobile app specifically designed to differentiate fatigue phenotypes using circadian symptom monitoring and state-of-the-art characterization of MS-related fatigue and its related symptoms. We also aim to report the first findings regarding patient compliance and the relationship between compliance and patient characteristics, including MS disease severity.

**Methods:** After developing the app, we used it in a prospective study designed to investigate the brain magnetic resonance imaging correlates of MS-related fatigue. In total, 64 patients with MS were recruited into this study and asked to use the app over a 2-week period. The app features the following modules: Visual Analogue Scales (VASs) to assess circadian changes in fatigue, depression, anxiety, and pain; daily sleep diaries (SLDs) to assess sleep habits and quality; and 10 one-time questionnaires to assess fatigue, depression, anxiety, sleepiness, physical activity, and motivation, as well as several other one-time questionnaires that were created to assess those relevant aspects of fatigue that were not captured by existing fatigue questionnaires. The app prompts subjects to assess their symptoms multiple times a day and enables real-time symptom monitoring through a web-accessible portal.

**Results:** Of 64 patients, 56 (88%) used the app, of which 51 (91%) completed all one-time questionnaires and 47 (84%) completed all one-time questionnaires, VASs, and SLDs. Patients reported no issues with the usage of the app, and there were no technical issues with our web-based data collection system. The relapsing-remitting MS to secondary-progressive MS ratio was significantly higher in patients who completed all one-time questionnaires, VASs, and SLDs than in those who completed all one-time questionnaires but not all VASs and SLDs ( $P=.01$ ). No other significant differences in demographics, fatigue, or disease severity were observed between the degrees of compliance.

**Conclusions:** The app can be used with reasonable compliance across patients with relapsing-remitting and secondary-progressive MS irrespective of demographics, fatigue, or disease severity.

(*JMIR Mhealth Uhealth* 2021;9(4):e19564) doi:[10.2196/19564](https://doi.org/10.2196/19564)

**KEYWORDS**

multiple sclerosis; fatigue; depression; mobile application; mobile phone; real-time assessment

## Introduction

### Background

Multiple sclerosis (MS) is an inflammatory, demyelinating disorder of the central nervous system that currently affects nearly 750,000 people in the United States [1]. Fatigue is one of the most disabling symptoms, affecting more than 65% of patients [2]. Fatigue has been associated with disease progression [3], and its pharmacological management remains challenging [4,5].

The pathogenesis of MS-related fatigue is not well understood. Neurogenic, inflammatory, endocrine, and metabolic mechanisms have been proposed [6]. In addition, several other factors, such as comorbid depression, anxiety, and sleep abnormalities, as well as physical activity and medications, may interfere with the perceived level of fatigue [2,6]. In fact, different fatigue phenotypes may exist in different patients. The ability to distinguish patient phenotypes might impact the understanding of underlying mechanisms and enable personalized management of this debilitating symptom. As an example of the need to capture the circadian nature of mood symptoms, including fatigue, it has been suggested that most people may experience lower energy levels in the evening than in the morning [7], whereas in a subset of patients affected by major depressive disorder, mood symptoms might be worse in the morning than in the evening [8]. In addition, in a recent study, we demonstrated the advantages of classifying patients according to temporal patterns of fatigue when associating damage to select brain circuitries with fatigue in patients with MS [9-14]. Although these studies highlighted the relevance of assessing temporal patterns of fatigue, retrospective data only included long-interval (every 1-2 years), repeated measures of fatigue, which might not reflect pathophysiologically relevant patterns.

### Objectives

We developed a mobile app to enable circadian assessment of fatigue and other mood symptoms, with the longer-term goal of identifying clinically and pathophysiologically relevant phenotypes of fatigue. As an example of its potential significance, this mobile app will enable us to test the hypothesis that diverse fatigue phenotypes may respond to different mechanism-specific treatments. Although several drugs have demonstrated efficacy in improving wakefulness in other conditions (such as narcolepsy [15]), none have been proven effective in treating fatigue in MS [4,5].

Several tools have been developed to assess various symptoms (including fatigue), monitor clinical status [14,16-20], or improve the perceived level of fatigue in patients with MS [16,19,21-23]. We believe that the presented mobile app meets a specific need for a tool to characterize fatigue phenotypes in

MS by assessing the temporal patterns of fatigue and its comorbid mood symptoms.

In this study, we aim to test and use this mobile app in the framework of a prospective study that was designed to investigate the association between MS-related fatigue and structural brain damage. The overarching goal of this study is to identify brain magnetic resonance imaging (MRI) predictors of persistent and treatment-resistant fatigue in MS. The aim of this study is to describe the design of the mobile app and to report the first findings regarding patient compliance and the relationship between compliance and patient characteristics, including MS disease severity.

## Methods

### Study Population and Study Design

Patients with MS were recruited from the Comprehensive Longitudinal Investigation of Multiple Sclerosis at Brigham and Women's Hospital (CLIMB) study [24] (n>2400), a large-scale, long-term study of patients with MS, using the following selection criteria: (1) availability of at least one previously recorded Modified Fatigue Impact Scale (MFIS) [25].

r Quality of Life in Neurological Disorders (Neuro-QoL) [26] score, (2) brain MRI scan acquired within 1 month of recruitment into the study, and (3) absence of clinical exacerbation within 3 months before their MRI. Selected patients were mailed a recruitment letter at least 2 weeks before their scheduled clinical visit at Brigham and Women's Hospital (BWH). The letter contained a brief description of our study procedures and invited patients who were interested in participating in our research study for an in-person study initiation session (SIS) following their scheduled clinical appointment or at another scheduled time at BWH. The aim of the SIS was to (1) provide patients with a detailed description of our study; (2) obtain a written informed consent form; and (3) teach patients how to use the following 3 study devices: the abovementioned mobile app installed on a mobile device, a wrist-worn actigraphic MotionLogger watch [27], and a Nox T3 home sleep test (HST) device [28]. The actigraphic watch assessed physical activity during the daytime and sleep quality at night throughout the entire study (ie, for 2 weeks), whereas the HST device was used to assess sleep apnea and periodic limb movements at one night in the patient's home. Data collected using the actigraphic watch and the HST device are not presented in this paper. Between May 2018 and September 2019, 64 patients with MS were recruited into the study and provided written informed consent in accordance with our study protocol approved by the institutional review board of BWH. Demographic data of these patients are presented in Table 1. Participants who answered all questions and returned all study devices to our laboratory received a remuneration of US \$100.



**Table 1.** Characteristics and compliance of the study participants.

Variable	All patients	Group 1: completed all one-time questionnaires, VASs <sup>a</sup> , and SLDs <sup>b</sup>	Group 2: completed all one-time questionnaires but not all VASs and SLDs	Group 3: answered some one-time questionnaires but no VAS or SLD	Group 4: did not answer any one-time questionnaire, VAS, or SLD
Subjects, n (%)	64 (100)	47 (73)	4 (6)	5 (8)	8 (13)
Age (years), mean (SD)	52 (9)	52 (8)	54.4 (10)	44.6 (10)	55.1 (8)
Sex (female), n (%)	54 (84)	40 (85)	4 (100)	5 (100)	4 (50)
Non-White or Hispanic, n (%)	3 (5)	2 (4)	0 (0)	1 (20)	0 (0)
<b>Disease category, n (%)</b>					
Relapsing-remitting multiple sclerosis	58 (100)	46 (79)	2 (3)	3 (5)	7 (12)
Secondary-progressive multiple sclerosis	5 (100)	1 (20)	2 (40)	1 (20)	1 (20)
Clinically isolated syndrome	1 (100)	0 (0)	0 (0)	1 (100)	0 (0)
Disease duration (years), mean (SD)	20 (8)	21 (8)	28 (10)	14 (6)	19 (7)
Expanded Disability Status Scale (years), mean (SD)	2.1 (1.5)	2.0 (1.1)	3.0 (2.4)	2.2 (2.5)	2.7 (2.3)
Fatigue Severity Scale score, mean (SD)	42 (13) <sup>c</sup>	41 (13)	47 (12)	43 (15) <sup>d</sup>	N/A <sup>e</sup>

<sup>a</sup>VASs: Visual Analogue Scales.

<sup>b</sup>SLDs: sleep diaries.

<sup>c</sup>In total, 55 patients completed the Fatigue Severity Scale. Therefore, the mean (SD) score of the Fatigue Severity Scale was calculated using 55 patients' data.

<sup>d</sup>One patient did not complete the Fatigue Severity Scale in this group. Therefore, the mean (SD) score of the Fatigue Severity Scale in this group was calculated using 4 patients' data.

<sup>e</sup>N/A: not applicable.

## Design of the Mobile App

Our research team, comprising physician scientists, designed and developed a mobile app in collaboration with the software developer team of *Mobilengine* [29]. Our teams communicated through virtual workshops. The app was installed on an Android smartphone with a 5-inch high-definition display that was provided to the participants. The participants used the smartphone for 2 weeks, during which they followed their normal daily routine. They then returned their smartphones to our research laboratory. The mobile app features 3 modules, which we describe in detail below, and in [Multimedia Appendix 1](#): (1) a series of one-time questionnaires, to be answered within

3 days of enrollment ([Table 2](#)); (2) Visual Analogue Scales (VASs) for self-reporting of fatigue, depression, anxiety, and pain levels, every 4 hours while awake; (3) a sleep diary (SLD) with separate items to be completed before and after each sleep or nap, containing a series of questions regarding perceived duration and quality of sleep, as well as the same VAS described earlier. VAS and SLD data were consistently collected for 14 days, starting after all one-time questionnaires were answered. Therefore, the duration of the data collection varied between 14 and 17 days depending on how much time it took for the participant to answer all one-time questionnaires. Of note, subjects were allowed to pause their entries and return to complete the remainder of the questions at a later time.

**Table 2.** Summary of existing questionnaires incorporated into our one-time questionnaire module.

Questionnaire	Domain assessed	Study
The Fatigue Severity Scale	Fatigue	Krupp et al (1989) [30]
The Neuro-QoL <sup>a</sup> fatigue questionnaire	Fatigue	Cella et al (2012) [26]
The Neuro-QoL depression questionnaire	Depression	Cella et al (2012) [26]
The Neuro-QoL anxiety questionnaire	Anxiety	Cella et al (2012) [26]
The Neuro-QoL sleep questionnaire	Sleep quality	Cella et al (2012) [26]
The Modified Fatigue Impact Scale	Fatigue	Amtmann et al (2012) [25]
Symptoms of depression questionnaire (7 questions only)	Vegetative symptoms of depression	Pedrelli et al (2014) [31]
The Epworth Sleepiness Scale	Sleepiness	Johns (1991) [32]
The Godin Leisure-Time Exercise questionnaire	Physical activity	Godin (1985) [33]
The Behavioral Approach System and Behavioral Avoidance System scale	Drive; fun seeking; reward responsiveness	Carver and White (1994) [34]

<sup>a</sup>Neuro-QoL: Quality of Life in Neurological Disorders.

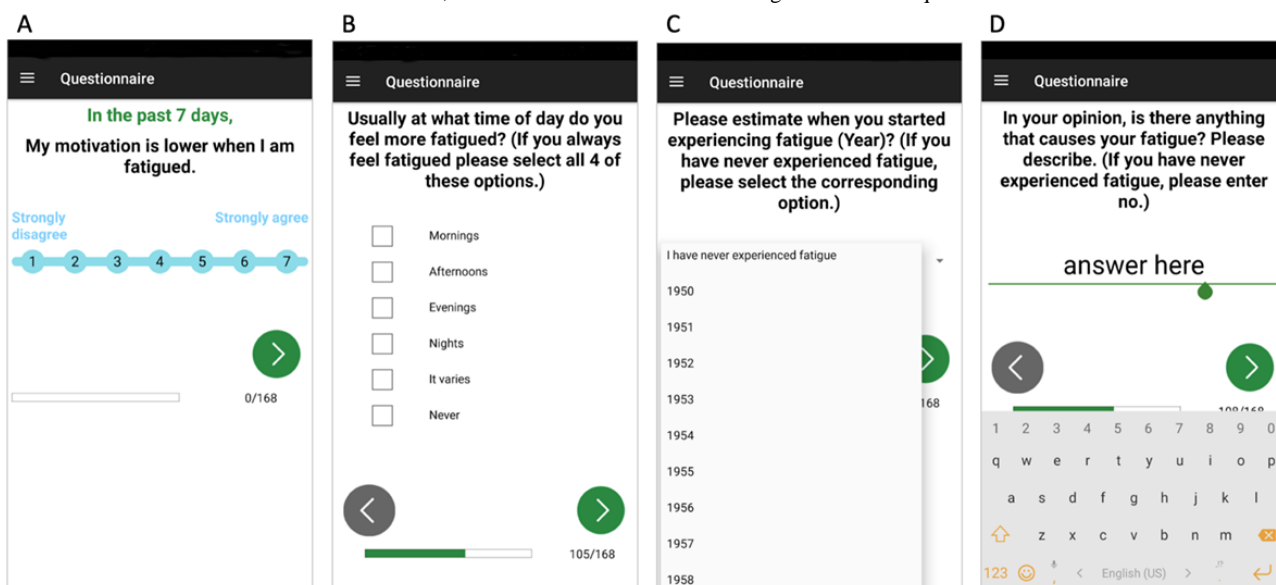
The following measures were taken to prevent slip errors: (1) user control and freedom: the selected (or entered) answer was submitted only when the user tapped the forward (or backward) arrow on the screen; (2) flexibility: users can freely navigate forward and backward between questions within the same block (ie, the same questionnaire) and make corrections as needed; (3) consistent and standard user interface: every scale had the same size and all questions, answer options, and forward and backward arrows were located on the same part of the screen. In addition, the app also has a landscape mode, in which the distance between the points of the VAS is larger; and (4) visibility and system status: the number of completed questions out of the total number of questions was shown at the bottom of the screen.

### ***One-Time Questionnaire Module***

We included 10 questionnaires (Table 2) that were previously validated. In addition, we created 22 new questions to assess the relevant aspects of fatigue that were not captured by existing

fatigue questionnaires, such as (1) circadian differences in fatigue severity and (2) effects of caffeine and nicotine use on the perceived level of fatigue, and (3) we incorporated open-ended questions aimed at exploring other potential aggravators and alleviators of fatigue (questions are presented on pages 16-20 of [Multimedia Appendix 1](#)). The one-time questionnaire module consisted of 168 questions. We preserved wording and the type of response (single choice, multiple choice, free text, analog scale, and selection of date) applied in the original questionnaires for each question included in our app (Figure 1). Questions were grouped into separate *blocks*, reflecting the original questionnaires. Patients were presented 1 question at a time but were able to freely navigate forward and backward between questions within the same block. Participants were notified at the end of each block and cautioned to be confident with their previous responses before final submission because they would no longer be able to return to a previous block of questions.

**Figure 1.** Sample screens for simple choice (A), multiple choice (B), number box (C), and text box (D) questions used in the one-time questionnaire module. Users can freely navigate forward and backward between questions within the same block using the forward (green) and backward (gray) arrows at the bottom of the screen. Under the arrows, a status bar indicates how far through the one-time questions a user is.

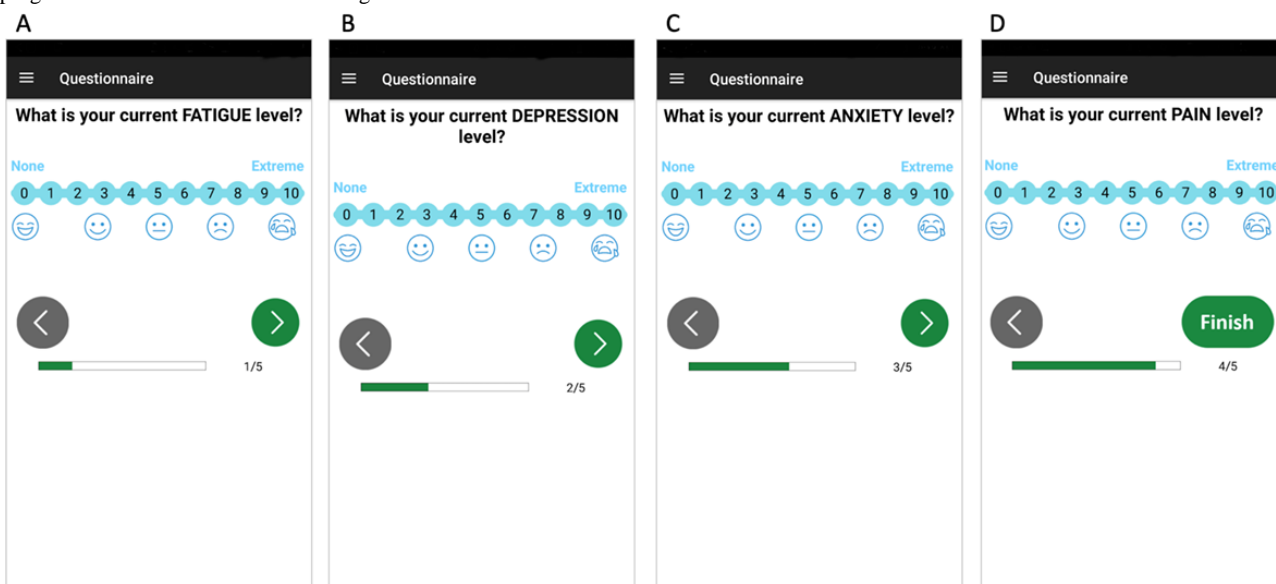


**VAS Module**

The VAS [35-39] module was used to assess a subject’s current level of fatigue, anxiety, depression, and pain. Each of these 4 VASs is scored between 0 and 10, with 0 representing absence (*none*) of the symptom (eg, *no fatigue*) and 10 representing *extreme* presence of the symptom (Figure 2). Previous studies have shown a very high correlation between VAS and a series of line drawings of faces with expressions of increasing distress [40,41]. We added a *faces* scale under each VAS under the assumption that the combination of the 2 scales may yield more robust results. Reminder functions were implemented in the mobile app based on the results of previous studies that associated reminders with positive effects on engagement with digital behavior change interventions [42]. One of our main aims was to assess circadian changes in fatigue, anxiety,

depression, and pain by measuring the level of these symptoms at least once in the morning, once in the afternoon, and once in the evening or night. Therefore, while awake, patients were prompted every 4 hours to complete the 4 VASs. Prompting was achieved through an acoustic and haptic (vibratory) alarm as well as a visual notification. In the absence of a response, the alarm and notification would be presented again after 30, 60, 90, and 120 minutes. If a subject did not answer a VAS within the allotted 2-hour period (between 4 and 6 hours after the previous VAS was completed), the subject would not be able to complete the VAS since beyond the allotted 2-hour period, the pending measurement would be closer in time to the next scheduled VAS assessment. Rather, they would be presented with notifications for a new VAS during an additional 2 hours (8 hours after the previous VAS had been completed).

**Figure 2.** Sample screens for Visual Analogue Scales for fatigue (A), depression (B), anxiety (C), and pain (D). Users can freely navigate forward and backward between questions using the forward (green) and backward (gray) arrows at the bottom of the screen. Under the arrows, a status bar indicates the progress of the user in the Visual Analogue Scale.

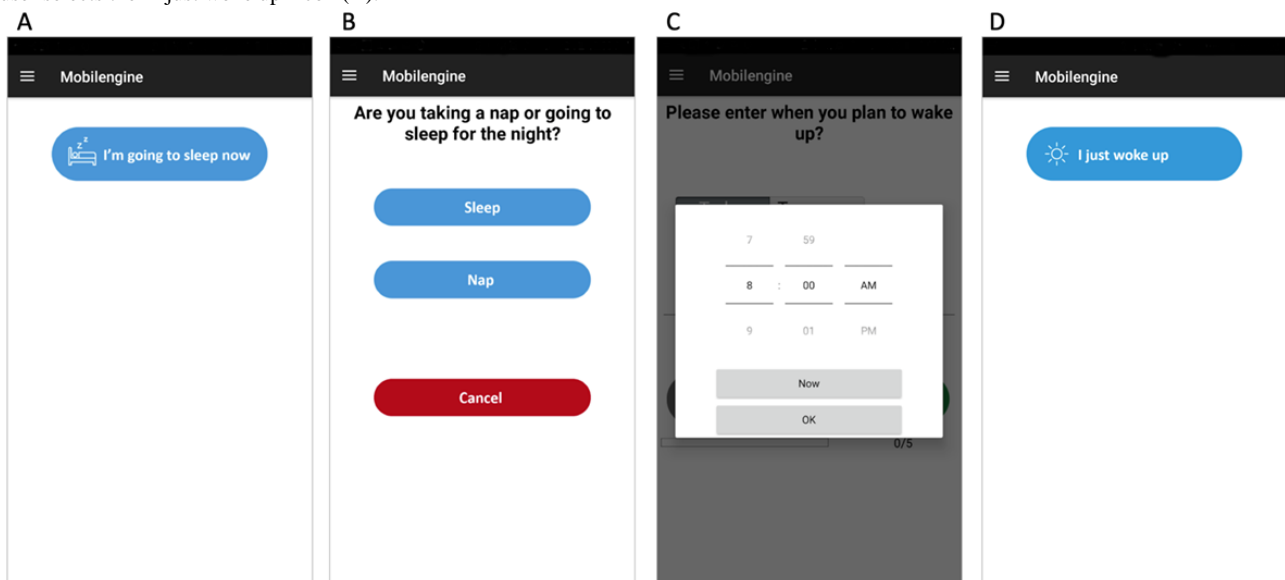


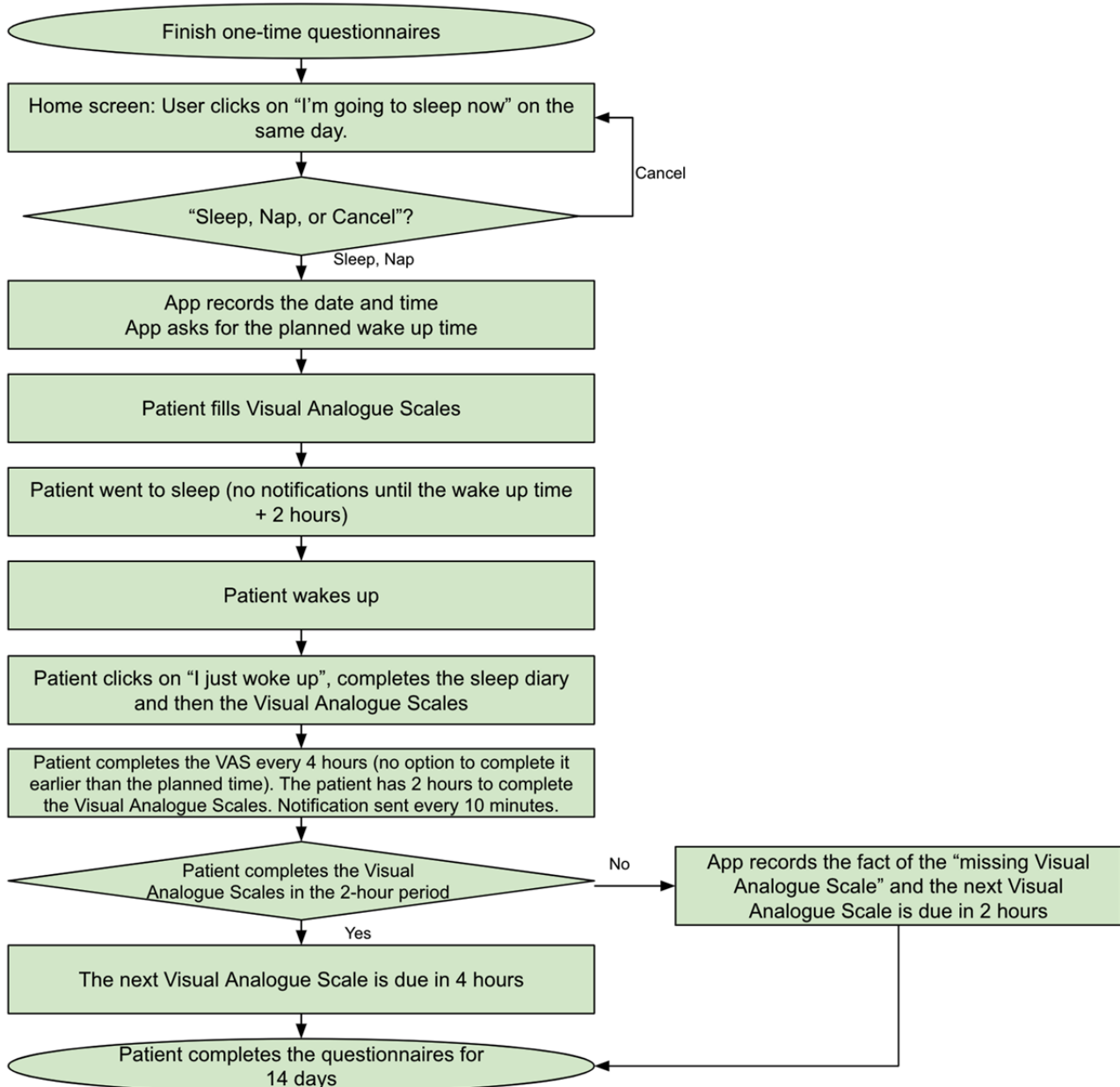
### SLD Module

The SLD contains items to be completed before sleep and items required after awakening from night sleep or daytime naps. At the time of recruitment, patients were instructed to activate the SLD immediately before any planned sleep period. Before sleep, the patient was prompted to indicate whether they were taking a nap or going to sleep for the night and their intended wake-up time (Figure 3). Then, the 4 VASs were presented to the patient for completion before the sleep episode. Sound alarms and notifications were deactivated in the mobile app until 2 hours following the intended wake-up time provided by the patient. Patients were at liberty to wake up spontaneously or prompted by a preprogrammed device, such as an alarm clock. On awakening, the patient was required to select the “I just woke up” icon on their phone (Figure 3). If this option was not activated 2 hours after the intended wake-up time, sound, haptic,

and visual alerts were activated. The “I just woke up” icon gives access to the wake items of the SLD that were developed by the Division of Sleep Medicine at BWH [43] and assesses final wake time, time awake before scheduled rising time, sleep latency, number of awakenings, sleep duration, subjective sleep quality, as well as current feeling of refreshed, sleepiness, and tenseness (wake items of the SLD are presented on page 21 of Multimedia Appendix 1). These items were followed by the day’s first VAS. Figure 4 shows the algorithm for the SLD and VAS assessments. If no sleep or nap option was selected on a given day, subjects were prompted by the app at 4 AM of the following calendar day, “We noticed that you did not indicate when you went to sleep last night. Please select the appropriate answer below:” with options for “I forgot to indicate when I went to sleep” and “I did not go to sleep.” If “I forgot to indicate when I went to sleep” was selected, the participant was prompted to enter an answer for “Time into bed” and to fill out an SLD.

**Figure 3.** Sample screens for the sleep diary module. Sleep diary is activated by tapping on the “I’m going to sleep now” icon (A). Then, the user is prompted to indicate whether they were taking a nap or going to sleep for the night (B), as well as their intended wake-up time (C). Upon awakening, the user selects the “I just woke up” icon (D).



**Figure 4.** Algorithm of Visual Analogue Scale and sleep diary assessments. VAS: Visual Analogue Scale.

The sequence of the questionnaires and questions was the same every time. All questions had to be answered before the questionnaire was submitted. At the bottom of the screen, a status bar indicated progress of the participant in the one-time questionnaires or VASs (Figures 1 and 2). The opening and submission dates and times were recorded for each question. All participants received an anonymized unique subject ID. No other identifying information was entered into the app.

### Monitoring the Database

All functions of the app were available offline (ie, with no internet connection) without limitations. The app saved and stored all subject responses locally in the smartphone's internal storage space. Each smartphone was provided with Wi-Fi and 3G (Third Generation) internet access to upload the collected data to Mobilengine's designated server as soon as each item was completed. If the network connection was inconsistent, the data were transferred as soon as the connection was adequate.

This system allowed us to monitor data collection in real time on demand. In addition, an email was sent daily at 4 PM Eastern time to designated investigators, listing the currently active devices with their IDs, as well as either (1) the progress of the subject in the one-time questionnaires or (2) whether the subject had filled out at least one VAS or SLD during the previous 24 hours. Subjects who did not complete their daily tasks are highlighted in red.

Following the completion of 14 days of VASs and SLDs, the subject's monitoring period was concluded, and they no longer had the option to answer further questions. The subject code was removed from daily emails. The ability to monitor our subjects' responses in this way allowed the investigators to promptly notify subjects who were not completing the study in the recommended time frame rather than waiting for the mobile device to be returned at the end of a 14- to 17-day period. We consider this monitoring feature to be an important part of mobile apps. All results in the database were easily exportable

as a table with rows containing subject ID, individual question ID, individual question response, and time of response. Every answer or item created a new row in the table (ie, a long data format).

### Assessment of Patient Compliance

Patient compliance was defined as adherence to the study protocol and was measured by the number of completed questions. We aimed to assess the association between the number of completed questions and demographic and clinical data. To this end, we grouped the recruited patients into the following 4 *compliance* groups: group 1: completed all one-time questionnaires, VASs, and SLDs; group 2: completed all one-time questionnaires but did not answer all VASs and SLDs; group 3: answered some, but not all one-time questionnaires and did not answer any VAS or SLD; group 4: did not answer any one-time questionnaire, VAS, or SLD. These groups were defined after the completion of patient recruitment. The following variables were compared between the 4 groups: age, sex, ethnicity, disease category (ie, relapsing-remitting multiple sclerosis [RRMS] and secondary-progressive multiple sclerosis [SPMS] ratio), disease duration, and physical disability (assessed using the Expanded Disability Status Scale [44]). These variables were assessed as part of the patients' routine clinical visits. We also compared the Fatigue Severity Scale (FSS) [30]

score between the abovementioned groups 1 to 3 (group 4 was not included in this analysis because patients in group 4 did not complete any questions) to assess whether baseline fatigue level was associated with patient compliance.

### Statistical Analysis

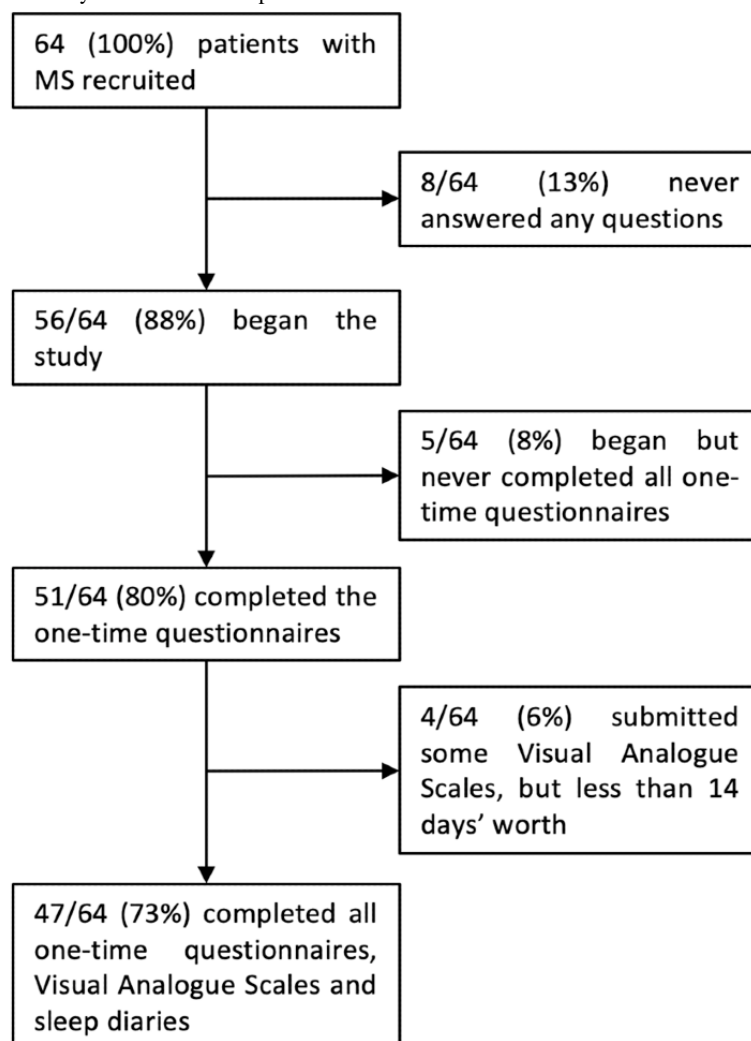
Differences in demographic and clinical variables between the patient compliance groups were assessed using analysis of variance or Kruskal-Wallis tests for continuous variables (depending on the distribution of the data) and the chi-square test or Fisher exact test (when  $n < 5$ ) for categorical variables. The threshold for statistical significance was set at  $P < .05$ . All statistical analyses were performed using Stata Statistical Software: Release 13 (StataCorp).

## Results

### Patient Compliance

Of the 64 patients, 56 (88%) began the study (ie, answered at least one question), whereas 8 (13%) patients did not complete any questions (Figure 5). Of the 56 patients who began the study, 51 (91%) completed the one-time questionnaires and 47 (84%) completed both the one-time questionnaire as well as the 14-day VAS and SLD modules (Figure 5).

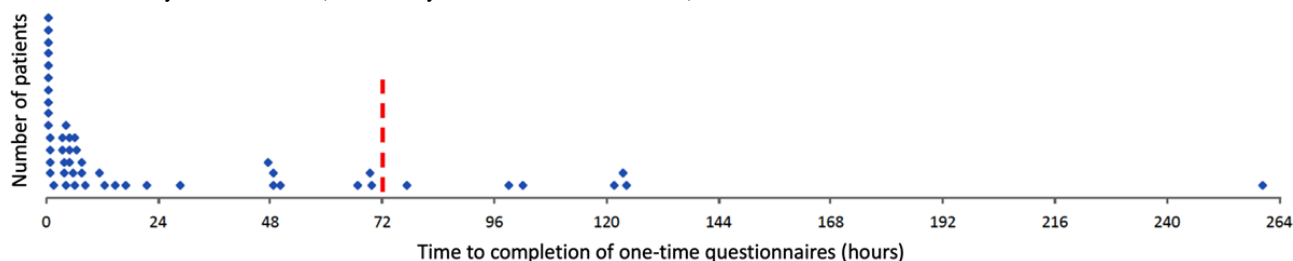
**Figure 5.** Patient compliance in our study cohort. MS: multiple sclerosis.



Among the 51 out of 64 patients who completed the one-time questionnaires, 44 (86%) met the target of 72 hours, and the median time to completion was 6.19 hours, with a range from 21 minutes to 10.9 days (Figure 3). Of the 51 patients, 14 (28%)

completed the one-time questionnaires within an hour, 20 additional patients (39%) completed it on the same day that they started it, and 5 additional patients (10%) completed the questionnaire within a 48-hour period (Figure 6).

**Figure 6.** Time to completion of one-time questionnaires (expressed in hours). Patients (indicated by blue rhombuses) were asked to answer all one-time questions within 3 days of enrollment (indicated by the dashed vertical red line).



All subjects who completed the one-time questionnaire module also completed at least 2 days of VAS modules. Of the 51 subjects, 3 (6%) stopped after 8 days or less.

Study subjects submitted close to 4 VASs per day with a mean of 3.9 (SD 1.3).

### Demographic and Clinical Comparison of Compliance Groups

There were no significant differences in age, sex, ethnicity, disease duration, and Expanded Disability Status Scale scores between the 4 patient compliance groups (Table 1). The RRMS/SPMS ratio was significantly higher in group 1 than in group 2 ( $P=.01$ ; Table 1). There was no significant difference in the RRMS/SPMS ratios between the other groups. There was no significant difference in FSS scores between groups 1 to 3.

## Discussion

### Principal Findings

We developed and used a mobile app that is innovative because of the following aspects: (1) assessment of fatigue and its associated mood symptoms (ie, depression and anxiety) and pain using the VAS every 4 hours while patients are awake for 2 weeks; (2) patient prompting to assess their symptoms multiple times a day; (3) real-time symptom monitoring of patients with MS by researchers or treating physicians through a web-accessible portal; and (4) single time point assessments of fatigue and its significant confounders, such as depression, anxiety, physical activity, sleep problems, and motivation level by commonly used questionnaires along with several other questions that we created to assess the relevant aspects of fatigue that were not captured by existing fatigue questionnaires. Here, we report good patient compliance with our mobile app-based assessments, that is, 91% (51/56) completed all one-time questionnaires and 84% (47/56) completed all one-time questionnaires, VASs, and SLDs out of those patients who started using the app ( $n=56$ ). Patients reported no issues with the usage of the app, and there were no technical issues with our web-based data collection system.

### Limitations

The limitations of the study are as follows: (1) we did not systematically collect feedback about the app itself, its ease of

use, or any patient suggestions about its features. However, our coinvestigators asked patients regarding their user experience and satisfaction when they contacted us (eg, because of the questions on the study protocol or the use of the study devices) or when our coinvestigators contacted the patients (eg, when the daily, system-generated email showed that a patient did not complete all questions on that day). Our high completion rate (47/56, 84% of those who began the study completed it) shows that many found it navigable and usable for the 2-week monitoring period. (2) The data collection and monitoring period was only 2 weeks. The observed patient compliance may change if the mobile app is used for a longer period (ie, beyond the novelty period) [45]. (3) The mobile app was installed on a smartphone that was provided to each participant. The use of a new device might have influenced the participants' behavior. (4) The current version of our app was not gamified. We may consider using gamification strategies when developing the next version of the app to further increase patient compliance and improve user experience.

### Comparison With Previous Work

Recently published studies underlined the relevance of delivery of health care through mobile devices (ie, mobile health) in MS [46,47]. Although several tools have been developed to assess various symptoms (including fatigue) and monitor clinical status [14,16-20,48] or to improve the perceived level of fatigue in MS [16,19,21-23], our app is the first that was specifically designed to conduct a state-of-the-art characterization of MS-related fatigue, including high-frequency fatigue assessments and real-time symptom monitoring.

To assess fatigue, 3 different self-assessment questionnaires (ie, MFIS [25], FSS [30], and Neuro-QoL fatigue [26]) and repeated daily VAS were implemented in our app. The MFIS has cognitive, physical, and psychosocial domains and addresses the effect of fatigue on daily activities [49]. The FSS addresses mainly physical fatigue with only one question related to cognitive fatigue [30]. The add-on value of the Neuro-QoL fatigue questionnaire is that it uses semantics to better characterize fatigue (eg, "felt exhausted," "had no energy," "felt fatigued," or "felt tired"). Both FSS and Neuro-QoL assess fatigue in the past week, whereas MFIS refers to fatigue in the past 4 weeks. In addition, we formulated several new questions to assess the relevant aspects of fatigue that were not captured

by existing fatigue questionnaires (eg, circadian differences in fatigue severity as well as potential aggravators and alleviators of fatigue).

Mood symptoms (ie, depression and anxiety) are highly intercorrelated with fatigue in MS [50,51], and their prevalence is over 20% in MS [52]. However, the relationship between MS-related fatigue and depression is not well understood [6,53]. Clinical symptomatology and, accordingly, clinical assessments or diagnosis of fatigue and depression show considerable overlap, as exemplified by (1) *fatigue or loss of interest* and *diminished ability to think or concentrate, or indecisiveness* are part of the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, diagnostic criteria of major depression; (2) *tiredness or fatigue, loss of energy*, and *concentration difficulty* are items on the Beck Depression Inventory [54]; and (3) impairment in concentration, thinking, and decision making are items in the MFIS [25]. Overlap in the clinical symptomatology of fatigue and depression may raise the question of whether these symptoms reflect separate entities with different etiologies and pathophysiologies or share a similar pathogenesis. To investigate the association between fatigue and depression and anxiety, we included two one-time questionnaires (ie, Neuro-QoL depression and anxiety batteries [26]) in our app along with daily VAS for depression and anxiety. In addition, we included those questions from the Symptoms of Depression questionnaire [31], which specifically assesses vegetative symptoms of depression under the hypothesis that the presence of vegetative symptoms distinguishes clinically significant depression from phenotypes of fatigue that are not associated with depression.

Other relevant confounders of fatigue include sleep abnormalities [6], physical activity [55], and motivation [56]. Patients with MS have significantly more sleep disturbances than the general population [6]. The prevalence of sleep abnormalities is approximately 50% in patients with MS [6]. To assess sleepiness and alterations in sleep quality, a one-time questionnaire (ie, Epworth sleepiness scale [32]) and daily SLD were incorporated into our mobile app.

Patients with MS may have a sedentary lifestyle because of associated disabilities (eg, muscle deconditioning or pain), diminished physical endurance, and deterioration of symptoms that could follow physical exertion [55]. However, a growing body of evidence suggests that nonpharmacological approaches (eg, exercise or psychobehavioral interventions) might be more

effective in reducing MS-related fatigue than medications [57]. We included a one-time questionnaire (ie, Godin Leisure-Time Exercise Questionnaire [33]) to assess physical activity and created targeted questions to assess the effect of physical activity, occupational, and psychological therapies on the patients' perceived level of fatigue.

The central fatigue model of Chaudhuri and Behan [58] hypothesized that central fatigue is caused by damage to the nonmotor components of the cortico-striato-thalamic system. Indeed, a more recent study associated MS-related fatigue with alterations in reward responsiveness [59]. Furthermore, several neuroimaging studies have found damage to reward-related brain regions in fatigued patients with MS [2,56]. Therefore, our app includes a self-assessment questionnaire (ie, behavioral approach system or behavioral avoidance system scale [34]) to measure reward responsiveness, drive, and fun seeking.

Previous studies have shown that fatigue fluctuates over time, even throughout the day [13,14]. Temporal fluctuations of fatigue are understudied in MS. Our recently published studies showed that taking into account temporal fatigue dynamics may improve the characterization of brain pathological correlates of MS-related fatigue in MS [9-11,60]. We included 4 VASs in our mobile app to assess circadian changes in fatigue, depression, anxiety, and pain levels under the hypothesis that the circadian evolution of these symptoms may show different patterns, potentially representing different fatigue phenotypes with different etiologies and pathogenesises.

To increase patient compliance, most of the abovementioned apps have built-in reminder functions, similar to our mobile app. In addition, our system provides real-time monitoring of data collection for researchers and patient-treating physicians. None of the abovementioned apps or platforms provided real-time monitoring of data collection.

## Conclusions

We developed a mobile app with reasonable patient compliance to assess fatigue and its confounders in patients with MS. This tool contains a battery of commonly used questionnaires and has the capability to monitor the level of subjective symptoms, such as fatigue, anxiety, depression, and pain, through VAS. We hope that it will facilitate remote monitoring of symptoms and adverse events and may allow more timely intervention than is possible with scheduled face-to-face visits.

## Acknowledgments

This investigation was supported by a grant from the National Multiple Sclerosis Society (grant RG-1501-03141) and Mobilengine. The processing fee of this manuscript was supported by a grant from Semmelweis University (grant EFOP-3.6.3-VEKOP-16-2017-00009).

## Conflicts of Interest

MP and MW reported no disclosures. GK is the director of the Mobilengine. AD is the CEO of the Mobilengine. CG has received support from Mobilengine (free use of platform and programming by Mobilengine Engineers), the National Multiple Sclerosis Society, the International Progressive Multiple Sclerosis Alliance, the US Office for Naval Research, and travel support from



Roche Pharmaceuticals; CG owns stock in Roche, Novartis, GlaxoSmithKline, Alnylam, Protalix Biotherapeutics, Arrowhead Pharmaceuticals, Cocrysal Pharma, and Sangamo Therapeutics.

### Multimedia Appendix 1

One-time questionnaire module: 11 previously validated questionnaires were implemented in the mobile app along with several additional questions that we formulated to assess relevant aspects of fatigue that were not captured by the existing questionnaires (eg, diurnal differences in fatigue severity, as well as potential aggravators and alleviators of fatigue).

[[PDF File \(Adobe PDF File\), 1411 KB - mhealth\\_v9i4e19564\\_app1.pdf](#)]

### References

1. Wallin MT, Culpepper WJ, Campbell JD, Nelson LM, Langer-Gould A, Marrie RA, et al. The prevalence of MS in the United States. *Neurology* 2019 Feb 15;92(10):1029-1040. [doi: [10.1212/wnl.0000000000007035](https://doi.org/10.1212/wnl.0000000000007035)]
2. Bakshi R. Fatigue associated with multiple sclerosis: diagnosis, impact and management. *Mult Scler* 2003 Jun;9(3):219-227. [doi: [10.1191/1352458503ms904oa](https://doi.org/10.1191/1352458503ms904oa)] [Medline: [12814166](https://pubmed.ncbi.nlm.nih.gov/12814166/)]
3. Cavallari M, Palotai M, Glanz BI, Egorova S, Prieto JC, Healy BC, et al. Fatigue predicts disease worsening in relapsing-remitting multiple sclerosis patients. *Mult Scler* 2016 Dec 11;22(14):1841-1849. [doi: [10.1177/1352458516635874](https://doi.org/10.1177/1352458516635874)] [Medline: [26920374](https://pubmed.ncbi.nlm.nih.gov/26920374/)]
4. Miller P, Soundy A. The pharmacological and non-pharmacological interventions for the management of fatigue related multiple sclerosis. *J Neurol Sci* 2017 Oct 15;381:41-54. [doi: [10.1016/j.jns.2017.08.012](https://doi.org/10.1016/j.jns.2017.08.012)] [Medline: [28991714](https://pubmed.ncbi.nlm.nih.gov/28991714/)]
5. Yang TT, Wang L, Deng XX, Yu G. Pharmacological treatments for fatigue in patients with multiple sclerosis: a systematic review and meta-analysis. *J Neurol Sci* 2017 Sep 15;380:256-261. [doi: [10.1016/j.jns.2017.07.042](https://doi.org/10.1016/j.jns.2017.07.042)] [Medline: [28870581](https://pubmed.ncbi.nlm.nih.gov/28870581/)]
6. Induruwa I, Constantinescu CS, Gran B. Fatigue in multiple sclerosis - a brief review. *J Neurol Sci* 2012 Dec 15;323(1-2):9-15. [doi: [10.1016/j.jns.2012.08.007](https://doi.org/10.1016/j.jns.2012.08.007)] [Medline: [22935407](https://pubmed.ncbi.nlm.nih.gov/22935407/)]
7. Wood C, Magnello ME. Diurnal changes in perceptions of energy and mood. *J R Soc Med* 1992;85(4):191-194 [FREE Full text] [Medline: [1433057](https://pubmed.ncbi.nlm.nih.gov/1433057/)]
8. Courtet P, Olié E. Circadian dimension and severity of depression. *Eur Neuropsychopharmacol* 2012 Jan;22 Suppl 3:476-481. [doi: [10.1016/j.euroneuro.2012.07.009](https://doi.org/10.1016/j.euroneuro.2012.07.009)] [Medline: [22959112](https://pubmed.ncbi.nlm.nih.gov/22959112/)]
9. Palotai M, Cavallari M, Healy BC, Guttman CR. A novel classification of fatigue in multiple sclerosis based on longitudinal assessments. *Mult Scler* 2020 May 23;26(6):725-734. [doi: [10.1177/1352458519898112](https://doi.org/10.1177/1352458519898112)] [Medline: [31971067](https://pubmed.ncbi.nlm.nih.gov/31971067/)]
10. Palotai M, Cavallari M, Koubiyar I, Pinzon AM, Nazeri A, Healy BC, et al. Microstructural fronto-striatal and temporo-insular alterations are associated with fatigue in patients with multiple sclerosis independent of white matter lesion load and depression. *Mult Scler* 2020 Nov 16;26(13):1708-1718. [doi: [10.1177/1352458519869185](https://doi.org/10.1177/1352458519869185)] [Medline: [31418637](https://pubmed.ncbi.nlm.nih.gov/31418637/)]
11. Palotai M, Nazeri A, Cavallari M, Healy BC, Glanz B, Gold SM, et al. History of fatigue in multiple sclerosis is associated with grey matter atrophy. *Sci Rep* 2019 Oct 14;9(1):14781 [FREE Full text] [doi: [10.1038/s41598-019-51110-2](https://doi.org/10.1038/s41598-019-51110-2)] [Medline: [31611598](https://pubmed.ncbi.nlm.nih.gov/31611598/)]
12. Palotai M, Guttman CR. Brain anatomical correlates of fatigue in multiple sclerosis. *Mult Scler* 2020 Jun 19;26(7):751-764. [doi: [10.1177/1352458519876032](https://doi.org/10.1177/1352458519876032)] [Medline: [31536461](https://pubmed.ncbi.nlm.nih.gov/31536461/)]
13. Johansson S, Ytterberg C, Hillert J, Widén Holmqvist L, von Koch L. A longitudinal study of variations in and predictors of fatigue in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2008 Apr 01;79(4):454-457. [doi: [10.1136/jnnp.2007.121129](https://doi.org/10.1136/jnnp.2007.121129)] [Medline: [18344396](https://pubmed.ncbi.nlm.nih.gov/18344396/)]
14. Powell DJH, Lioffi C, Schlotz W, Moss-Morris R. Tracking daily fatigue fluctuations in multiple sclerosis: ecological momentary assessment provides unique insights. *J Behav Med* 2017 Oct 9;40(5):772-783 [FREE Full text] [doi: [10.1007/s10865-017-9840-4](https://doi.org/10.1007/s10865-017-9840-4)] [Medline: [28281106](https://pubmed.ncbi.nlm.nih.gov/28281106/)]
15. Billiard M, Broughton R. Modafinil: its discovery, the early European and North American experience in the treatment of narcolepsy and idiopathic hypersomnia, and its subsequent use in other medical conditions. *Sleep Med* 2018 Sep;49:69-72. [doi: [10.1016/j.sleep.2018.05.027](https://doi.org/10.1016/j.sleep.2018.05.027)] [Medline: [30174215](https://pubmed.ncbi.nlm.nih.gov/30174215/)]
16. Alexander S, Peryer G, Gray E, Barkhof F, Chataway J. Wearable technologies to measure clinical outcomes in multiple sclerosis: a scoping review. *Mult Scler* 2020 Aug 04:- [FREE Full text] [doi: [10.1177/1352458520946005](https://doi.org/10.1177/1352458520946005)] [Medline: [32749928](https://pubmed.ncbi.nlm.nih.gov/32749928/)]
17. Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures. *Neurol Neuroimmunol Neuroinflamm* 2015 Oct 15;2(6):e162. [doi: [10.1212/nxi.0000000000000162](https://doi.org/10.1212/nxi.0000000000000162)]
18. Newland P, Oliver B, Newland J, Thomas F. Testing feasibility of a mobile application to monitor fatigue in people with multiple sclerosis. *J Neurosci Nurs* 2019 Dec;51(6):331-334. [doi: [10.1097/JNN.0000000000000479](https://doi.org/10.1097/JNN.0000000000000479)] [Medline: [31688283](https://pubmed.ncbi.nlm.nih.gov/31688283/)]
19. Giunti G, Fernández EG, Zubieta ED, Rivera Romero O. Supply and demand in mhealth apps for persons with multiple sclerosis: systematic search in app stores and scoping literature review. *JMIR Mhealth Uhealth* 2018 May 23;6(5):- [FREE Full text] [doi: [10.2196/10512](https://doi.org/10.2196/10512)] [Medline: [29792295](https://pubmed.ncbi.nlm.nih.gov/29792295/)]
20. Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, et al. Evaluating the utility of smartphone-based sensor assessments in persons with multiple sclerosis in the real-world using an app (elevateMS): observational, prospective

- pilot digital health study. *JMIR Mhealth Uhealth* 2020 Oct 27;8(10):- [FREE Full text] [doi: [10.2196/22108](https://doi.org/10.2196/22108)] [Medline: [33107827](https://pubmed.ncbi.nlm.nih.gov/33107827/)]
21. Van Geel F, Geurts E, Abasiyanik Z, Coninx K, Feys P. Feasibility study of a 10-week community-based program using the WalkWithMe application on physical activity, walking, fatigue and cognition in persons with Multiple Sclerosis. *Mult Scler Relat Disord* 2020 Jul;42:-. [doi: [10.1016/j.msard.2020.102067](https://doi.org/10.1016/j.msard.2020.102067)] [Medline: [32371377](https://pubmed.ncbi.nlm.nih.gov/32371377/)]
  22. D'hooghe M, Van Gassen G, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: the MS TeleCoach feasibility study. *Mult Scler Relat Disord* 2018 May;22:90-96 [FREE Full text] [doi: [10.1016/j.msard.2018.03.020](https://doi.org/10.1016/j.msard.2018.03.020)] [Medline: [29649789](https://pubmed.ncbi.nlm.nih.gov/29649789/)]
  23. Giunti G, Mylonopoulou V, Romero OR. 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](https://doi.org/10.2196/mhealth.9437)] [Medline: [29500159](https://pubmed.ncbi.nlm.nih.gov/29500159/)]
  24. Gauthier SA, Glanz BI, Mandel M, Weiner HL. A model for the comprehensive investigation of a chronic autoimmune disease: the multiple sclerosis CLIMB study. *Autoimmun Rev* 2006 Oct;5(8):532-536. [doi: [10.1016/j.autrev.2006.02.012](https://doi.org/10.1016/j.autrev.2006.02.012)] [Medline: [17027888](https://pubmed.ncbi.nlm.nih.gov/17027888/)]
  25. Amtmann D, Bamer AM, Noonan V, Lang N, Kim J, Cook KF. Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. *Rehabil Psychol* 2012 May;57(2):159-166 [FREE Full text] [doi: [10.1037/a0027890](https://doi.org/10.1037/a0027890)] [Medline: [22686554](https://pubmed.ncbi.nlm.nih.gov/22686554/)]
  26. Cella D, Lai JS, Nowinski CJ, Victorson D, Peterman A, Miller D, et al. Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. *Neurology* 2012 May 09;78(23):1860-1867. [doi: [10.1212/wnl.0b013e318258f744](https://doi.org/10.1212/wnl.0b013e318258f744)]
  27. Actigraphy for sleep, chronobiology and physical activity. MotionWatch 8. URL: [https://www.camntech.com/motionwatch-8/?gclid=CjwKCAjw0On8BRAGeIwAincsHOyMAfBumqmYOVFYkOtvDsv-yJwFhYGcz9oH4dohxUc-8KACIRcxoCUHYQAvD\\_BwE](https://www.camntech.com/motionwatch-8/?gclid=CjwKCAjw0On8BRAGeIwAincsHOyMAfBumqmYOVFYkOtvDsv-yJwFhYGcz9oH4dohxUc-8KACIRcxoCUHYQAvD_BwE) [accessed 2021-03-12]
  28. Nox T3 HST. Nox Medical. URL: [https://noxmedical.com/products/nox-t3-sleep-monitor/?gclid=CjwKCAjwltH3BRB6EiwAhj0IUH7wGNafOtMFTFARWolsH-Ocga4CbvpAHWvyvSSE-4OzJLQ4EEMeXMhoC3GkQAvD\\_BwE](https://noxmedical.com/products/nox-t3-sleep-monitor/?gclid=CjwKCAjwltH3BRB6EiwAhj0IUH7wGNafOtMFTFARWolsH-Ocga4CbvpAHWvyvSSE-4OzJLQ4EEMeXMhoC3GkQAvD_BwE) [accessed 2021-03-12]
  29. Mobilengine. URL: <http://www.mobilengine.com/> [accessed 2021-03-12]
  30. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989 Oct 01;46(10):1121-1123. [doi: [10.1001/archneur.1989.00520460115022](https://doi.org/10.1001/archneur.1989.00520460115022)] [Medline: [2803071](https://pubmed.ncbi.nlm.nih.gov/2803071/)]
  31. Pedrelli P, Blais MA, Alpert JE, Shelton RC, Walker RSW, Fava M. Reliability and validity of the Symptoms of Depression Questionnaire (SDQ). *CNS Spectr* 2014 Oct 02;19(6):535-546. [doi: [10.1017/s1092852914000406](https://doi.org/10.1017/s1092852914000406)]
  32. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991 Dec;14(6):540-545. [doi: [10.1093/sleep/14.6.540](https://doi.org/10.1093/sleep/14.6.540)] [Medline: [1798888](https://pubmed.ncbi.nlm.nih.gov/1798888/)]
  33. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Can J Appl Sport Sci* 1985 Sep;10(3):141-146. [Medline: [4053261](https://pubmed.ncbi.nlm.nih.gov/4053261/)]
  34. Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS Scales. *J Pers Soc Psychol* 1994;67(2):319-333. [doi: [10.1037/0022-3514.67.2.319](https://doi.org/10.1037/0022-3514.67.2.319)]
  35. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain* 2001 Aug;93(2):173-183. [doi: [10.1016/S0304-3959\(01\)00314-1](https://doi.org/10.1016/S0304-3959(01)00314-1)] [Medline: [11427329](https://pubmed.ncbi.nlm.nih.gov/11427329/)]
  36. Téllez N, Ríó J, Tintoré M, Nos C, Galán I, Montalban X. Does the Modified Fatigue Impact Scale offer a more comprehensive assessment of fatigue in MS? *Mult Scler* 2005 Apr 02;11(2):198-202. [doi: [10.1191/1352458505ms1148oa](https://doi.org/10.1191/1352458505ms1148oa)] [Medline: [15794395](https://pubmed.ncbi.nlm.nih.gov/15794395/)]
  37. Ahearn EP. The use of visual analog scales in mood disorders: a critical review. *J Psychiatr Res* 1997 Sep;31(5):569-579. [doi: [10.1016/s0022-3956\(97\)00029-0](https://doi.org/10.1016/s0022-3956(97)00029-0)]
  38. Gur-Ozmen S, Leibetseder A, Cock HR, Agrawal N, von Oertzen TJ. Screening of anxiety and quality of life in people with epilepsy. *Seizure* 2017 Feb;45:107-113 [FREE Full text] [doi: [10.1016/j.seizure.2016.11.026](https://doi.org/10.1016/j.seizure.2016.11.026)] [Medline: [27984808](https://pubmed.ncbi.nlm.nih.gov/27984808/)]
  39. Kim SH, Kim YH, Kim HJ. Laughter and stress relief in cancer patients: a pilot study. *Evid Based Complement Alternat Med* 2015;2015:- [FREE Full text] [doi: [10.1155/2015/864739](https://doi.org/10.1155/2015/864739)] [Medline: [26064177](https://pubmed.ncbi.nlm.nih.gov/26064177/)]
  40. Fadaizadeh L, Emami H, Samii K. Comparison of visual analogue scale and faces rating scale in measuring acute postoperative pain. *Arch Iran Med* 2009 Jan;12(1):73-75 [FREE Full text] [Medline: [19111034](https://pubmed.ncbi.nlm.nih.gov/19111034/)]
  41. Malviya S, Polaner D, Berde C. Acute pain. In: *A Practice of Anesthesia for Infants and Children (Fourth Edition)*. Amsterdam: Elsevier; 2009:939-978.
  42. 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* 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]

43. O'Donnell D, Silva EJ, Münch M, Ronda JM, Wang W, Duffy JF. Comparison of subjective and objective assessments of sleep in healthy older subjects without sleep complaints. *J Sleep Res* 2009 Jun;18(2):254-263 [FREE Full text] [doi: [10.1111/j.1365-2869.2008.00719.x](https://doi.org/10.1111/j.1365-2869.2008.00719.x)] [Medline: [19645969](https://pubmed.ncbi.nlm.nih.gov/19645969/)]
44. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983 Nov 01;33(11):1444-1452. [doi: [10.1212/wnl.33.11.1444](https://doi.org/10.1212/wnl.33.11.1444)] [Medline: [6685237](https://pubmed.ncbi.nlm.nih.gov/6685237/)]
45. Shin G, Feng Y, Jarrahi MH, Gafinowitz N. Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. *JAMIA Open* 2019 Apr;2(1):62-72 [FREE Full text] [doi: [10.1093/jamiaopen/ooy048](https://doi.org/10.1093/jamiaopen/ooy048)] [Medline: [31984346](https://pubmed.ncbi.nlm.nih.gov/31984346/)]
46. Giunti G, Kool J, Romero OR, Zubiete ED. Exploring the specific needs of persons with multiple sclerosis for mhealth solutions for physical activity: mixed-methods study. *JMIR Mhealth Uhealth* 2018 Feb 09;6(2):e37 [FREE Full text] [doi: [10.2196/mhealth.8996](https://doi.org/10.2196/mhealth.8996)] [Medline: [29426814](https://pubmed.ncbi.nlm.nih.gov/29426814/)]
47. Van Kessel K, Babbage DR, Reay N, Miner-Williams WM, Kersten P. Mobile technology use by people experiencing multiple sclerosis fatigue: survey methodology. *JMIR Mhealth Uhealth* 2017 Feb 28;5(2):e6 [FREE Full text] [doi: [10.2196/mhealth.6192](https://doi.org/10.2196/mhealth.6192)] [Medline: [28246073](https://pubmed.ncbi.nlm.nih.gov/28246073/)]
48. Polhemus AM, Novák J, Ferrao J, Simblett S, Radaelli M, Locatelli P, et al. Human-centered design strategies for device selection in mhealth programs: development of a novel framework and case study. *JMIR Mhealth Uhealth* 2020 May 07;8(5):- [FREE Full text] [doi: [10.2196/16043](https://doi.org/10.2196/16043)] [Medline: [32379055](https://pubmed.ncbi.nlm.nih.gov/32379055/)]
49. Flachenecker P, Kümpfel T, Kallmann B, Gottschalk M, Grauer O, Rieckmann P, et al. Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. *Mult Scler* 2002 Dec 02;8(6):523-526. [doi: [10.1191/1352458502ms839oa](https://doi.org/10.1191/1352458502ms839oa)] [Medline: [12474995](https://pubmed.ncbi.nlm.nih.gov/12474995/)]
50. Palotai M, Mike A, Cavallari M, Strammer E, Orsi G, Healy BC, et al. Changes to the septo-fornical area might play a role in the pathogenesis of anxiety in multiple sclerosis. *Mult Scler* 2018 Jul 26;24(8):1105-1114. [doi: [10.1177/1352458517711273](https://doi.org/10.1177/1352458517711273)] [Medline: [28548605](https://pubmed.ncbi.nlm.nih.gov/28548605/)]
51. Ive P, MacLeod W, Mkumla N, Orrell C, Jentsch U, Wallis CL, et al. Low prevalence of liver disease but regional differences in HBV treatment characteristics mark HIV/HBV co-infection in a South African HIV clinical trial. *PLoS One* 2013 Dec 6;8(12):- [FREE Full text] [doi: [10.1371/journal.pone.0074900](https://doi.org/10.1371/journal.pone.0074900)] [Medline: [24324573](https://pubmed.ncbi.nlm.nih.gov/24324573/)]
52. Marrie RA, Reingold S, Cohen J, Stuve O, Trojano M, Sorensen PS, et al. The incidence and prevalence of psychiatric disorders in multiple sclerosis: a systematic review. *Mult Scler* 2015 Mar 12;21(3):305-317 [FREE Full text] [doi: [10.1177/1352458514564487](https://doi.org/10.1177/1352458514564487)] [Medline: [25583845](https://pubmed.ncbi.nlm.nih.gov/25583845/)]
53. Feinstein A, Magalhaes S, Richard JF, Audet B, Moore C. The link between multiple sclerosis and depression. *Nat Rev Neurol* 2014 Sep 12;10(9):507-517. [doi: [10.1038/nrneuro.2014.139](https://doi.org/10.1038/nrneuro.2014.139)] [Medline: [25112509](https://pubmed.ncbi.nlm.nih.gov/25112509/)]
54. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961 Jul;4:561-571. [doi: [10.1001/archpsyc.1961.01710120031004](https://doi.org/10.1001/archpsyc.1961.01710120031004)] [Medline: [13688369](https://pubmed.ncbi.nlm.nih.gov/13688369/)]
55. Chaudhuri A, Behan PO. Fatigue in neurological disorders. *The Lancet* 2004 Mar;363(9413):978-988. [doi: [10.1016/s0140-6736\(04\)15794-2](https://doi.org/10.1016/s0140-6736(04)15794-2)]
56. Dobryakova E, DeLuca J, Genova HM, Wylie GR. Neural correlates of cognitive fatigue: cortico-striatal circuitry and effort-reward imbalance. *J Int Neuropsychol Soc* 2013 Jul 10;19(8):849-853. [doi: [10.1017/s1355617713000684](https://doi.org/10.1017/s1355617713000684)]
57. Asano M, Finlayson ML. Meta-analysis of three different types of fatigue management interventions for people with multiple sclerosis: exercise, education, and medication. *Mult Scler Int* 2014;2014:--- [FREE Full text] [doi: [10.1155/2014/798285](https://doi.org/10.1155/2014/798285)] [Medline: [24963407](https://pubmed.ncbi.nlm.nih.gov/24963407/)]
58. Chaudhuri A, Behan PO. Fatigue and basal ganglia. *J Neurol Sci* 2000 Oct;179(1-2):34-42. [doi: [10.1016/s0022-510x\(00\)00411-1](https://doi.org/10.1016/s0022-510x(00)00411-1)]
59. Pardini M, Capello E, Krueger F, Mancardi G, Uccelli A. Reward responsiveness and fatigue in multiple sclerosis. *Mult Scler* 2013 Feb 21;19(2):233-240. [doi: [10.1177/1352458512451509](https://doi.org/10.1177/1352458512451509)] [Medline: [22723570](https://pubmed.ncbi.nlm.nih.gov/22723570/)]
60. Palotai M, Small C, Makris N, Somes NG, Pinzon AM, Rathi Y, et al. Microstructural changes in the left mesocorticolimbic pathway are associated with the comorbid development of fatigue and depression in multiple sclerosis. *J Neuroimaging* 2021 Mar 01:- epub ahead of print. [doi: [10.1111/jon.12832](https://doi.org/10.1111/jon.12832)] [Medline: [33522683](https://pubmed.ncbi.nlm.nih.gov/33522683/)]

## Abbreviations

**BWH:** Brigham and Women's Hospital

**CLIMB:** Comprehensive Longitudinal Investigation of Multiple Sclerosis at Brigham and Women's Hospital

**FSS:** Fatigue Severity Scale

**HST:** home sleep test

**MFIS:** Modified Fatigue Impact Scale

**MRI:** magnetic resonance imaging

**MS:** multiple sclerosis

**Neuro-QoL:** Quality of Life in Neurological Disorders

**RRMS:** relapsing-remitting multiple sclerosis

**SIS:** study initiation session

**SLD:** sleep diary

**SPMS:** secondary-progressive multiple sclerosis

**VAS:** Visual Analogue Scale

*Edited by L Buis; submitted 23.04.20; peer-reviewed by A Tarnoki, V Mylonopoulou, G Giunti; comments to author 04.09.20; revised version received 10.12.20; accepted 02.02.21; published 16.04.21.*

*Please cite as:*

*Palotai M, Wallack M, Kujbus G, Dalnoki A, Guttmann C*

*Usability of a Mobile App for Real-Time Assessment of Fatigue and Related Symptoms in Patients With Multiple Sclerosis: Observational Study*

*JMIR Mhealth Uhealth 2021;9(4):e19564*

*URL: <https://mhealth.jmir.org/2021/4/e19564>*

*doi: [10.2196/19564](https://doi.org/10.2196/19564)*

*PMID: [33861208](https://pubmed.ncbi.nlm.nih.gov/33861208/)*

©Miklos Palotai, Max Wallack, Gergo Kujbus, Adam Dalnoki, Charles Guttmann. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 16.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Slip Buddy App for Weight Management: Randomized Feasibility Trial of a Dietary Lapse Tracking App

Sherry Pagoto<sup>1</sup>, PhD; Bengisu Tulu<sup>2</sup>, PhD; Molly E Waring<sup>1</sup>, PhD; Jared Goetz<sup>1</sup>, BSc; Jessica Bibeau<sup>1</sup>, MA; Joseph Divito<sup>1</sup>, BA; Laurie Groshon<sup>1</sup>, BA; Matthew Schroeder<sup>1</sup>, BA

<sup>1</sup>University of Connecticut, Department of Allied Health Sciences, Storrs, CT, United States

<sup>2</sup>Worcester Polytechnic University, Foisie Business School, Worcester, MA, United States

**Corresponding Author:**

Sherry Pagoto, PhD

University of Connecticut

Department of Allied Health Sciences

2006 Hillside Road, Unit 1248, Room 22

Storrs, CT, 06269

United States

Phone: 1 5084864900

Email: [Sherry.Pagoto@uconn.edu](mailto:Sherry.Pagoto@uconn.edu)

## Abstract

**Background:** Although calorie tracking is one of the strongest predictors of weight loss in behavioral weight loss interventions, low rates of adherence are common.

**Objective:** This study aims to examine the feasibility and acceptability of using the Slip Buddy app during a 12-week web-based weight loss program.

**Methods:** We conducted a randomized pilot trial to evaluate the feasibility and acceptability of using the Slip Buddy app compared with a popular commercial calorie tracking app during a counselor-led, web-based behavioral weight loss intervention. Adults who were overweight or obese were recruited on the web and randomized into a 12-week web-based weight loss intervention that included either the Slip Buddy app or a commercial calorie tracking app. Feasibility outcomes included retention, app use, usability, slips reported, and contextual factors reported at slips. Acceptability outcomes included ratings of how helpful, tedious, taxing, time consuming, and burdensome using the assigned app was. We described weight change from baseline to 12 weeks in both groups as an exploratory outcome. Participants using the Slip Buddy app provided feedback on how to improve it during the postintervention focus groups.

**Results:** A total of 75% (48/64) of the participants were female and, on average, 39.8 (SD 11.0) years old with a mean BMI of 34.2 (SD 4.9) kg/m<sup>2</sup>. Retention was high in both conditions, with 97% (31/32) retained in the Slip Buddy condition and 94% (30/32) retained in the calorie tracking condition. On average, participants used the Slip Buddy app on 53.8% (SD 31.3%) of days, which was not significantly different from those using the calorie tracking app (mean 57.5%, SD 28.4% of days), and participants who recorded slips (30/32, 94%) logged on average 17.9 (SD 14.4) slips in 12 weeks. The most common slips occurred during snack times (220/538, 40.9%). Slips most often occurred at home (297/538, 55.2%), while working (153/538, 28.4%), while socializing (130/538, 24.2%), or during screen time (123/538, 22.9%). The conditions did not differ in participants' ratings of how their assigned app was tedious, taxing, or time consuming (all values of  $P > .05$ ), but the calorie tracking condition gave their app higher helpfulness and usability ratings (all values of  $P < .05$ ). Technical issues were the most common type of negative feedback, whereas simplicity was the most common type of positive feedback. Weight losses of  $\geq 5\%$  of baseline weight were achieved by 31% (10/32) of Slip Buddy participants and 34% (11/32) of calorie tracking participants.

**Conclusions:** Self-monitoring of dietary lapses and the contextual factors associated with them may be an alternative for people who do not prefer calorie tracking. Future research should examine patient characteristics associated with adherence to different forms of dietary self-monitoring.

**Trial Registration:** ClinicalTrials.gov NCT02615171; <https://clinicaltrials.gov/ct2/show/NCT02615171>

(*JMIR Mhealth Uhealth* 2021;9(4):e24249) doi:[10.2196/24249](https://doi.org/10.2196/24249)

**KEYWORDS**

mobile app; mHealth; weight loss; obesity; diet; mobile phone

## Introduction

### Background

Obesity is a major risk factor for type 2 diabetes [1], but many people still do not have access to evidence-based lifestyle interventions [2]. Technology-delivered lifestyle interventions may have increased reach, and early studies reveal promising impacts, but they are still fairly burdensome and expensive [3]. A key source of burden is self-monitoring in the form of calorie tracking, which requires a person to record all of the food and beverage they consume each day. Although adherence to calorie tracking predicts weight loss outcomes [4], adherence is notoriously low [5]. A recent study found that rates of consistent calorie tracking in a web-based weight loss program fell from 68% in week 1 to 21% by week 12 [6]. Interventions that do not use calorie tracking have failed to produce weight loss outcomes [7]. Thus, new forms of self-monitoring that can produce similar or greater weight loss are needed.

Interestingly, calorie tracking is a relatively complex form of self-monitoring compared with forms used for other behaviors. For example, in smoking cessation interventions, the smoker keeps a tally of the number of cigarettes smoked each day and notes the triggers associated with each smoking episode [8]. Similarly, a simpler way to perform dietary self-monitoring could be to have people track only the segment of their diet that requires intervention, that is, the eating episodes that account for excess calories or dietary lapses. Dietary lapses are considered *nonhomeostatic eating*, which includes any eating episode that occurs in excess of one's needs to maintain health, not just in terms of food quantity but also food quality [9]. To the extent that we can help people identify and eliminate dietary lapses, we may be able to affect energy balance without the task of calorie tracking.

Previous ecological momentary assessment studies have revealed that people are able to identify and self-monitor dietary lapses [10-13]. Carels et al [12] had 12 dieters track their dietary lapses over the course of a week along with contextual factors surrounding the lapses. The results showed that hunger, location, negative affect, and certain activities such as watching television and socializing were common during lapses. Stress, hunger, and socializing were found to be predictors of dietary lapses in other studies [13,14]. This research has stimulated at least two dietary lapse tracking apps designed for use in behavioral weight loss interventions [15,16]. OnTrack allows users to track their dietary lapses from the WW (formerly Weight Watchers) points-based weight loss program and the contextual factors surrounding those lapses. The app sends 6 prompts per day asking the users to record 17 contextual factors (eg, mood, hunger, and temptations), the data from which go into a machine learning algorithm that drives just-in-time intervention messages that occur when the user is at high risk for lapses [10,15]. In a randomized trial, adding OnTrack to WW's points-based weight loss program was found superior in weight loss outcomes over 10 weeks compared with the WW program alone, but it did not

appear to enhance weight loss when added to a version of the WW program that was less intensive in terms of point tracking [15]. In contrast to OnTrack, which requires users to perform WW point tracking, the Slip Buddy app was designed as a replacement for traditional forms of dietary tracking.

Employing a user-centered design, we developed the Slip Buddy app, which allows users to track dietary lapses as they occur and the contextual factors that cued each lapse [16]. To best mimic the way calorie tracking apps are used, users are only expected to record at the time of recordable eating episodes. However, unlike traditional calorie tracking apps that require the user to record all intake every day, the burden of using Slip Buddy declines as dieters become more skilled at avoiding lapses simply because the only task of the app is to track lapses. Slip Buddy works such that users simply hit an *Oops!* button each time they experience a dietary lapse, which are referred to as slips in the app based on feedback regarding preferred language from the target population during the design process. When a user logs a slip, the app asks them to rate stress and degree of hunger to satiety on scales of 0 to 10 and to name the location (eg, home and work) and activity (eg, watching television and socializing) they were engaging in at the time of the slip and the food consumed. The app passively collects day and time information and maintains a log of all slips. A tab called *Slip History* shows the user's entire history of slips, including contextual data (eg, stress rating and location) so that users can learn about the circumstances that trigger their slips. Each morning, the user receives a notification to complete a daily check-in that involves recording how long they slept, their stress and hunger level, and their weight that morning if they stepped on the scale. Each afternoon, they receive a notification to complete check-in on their stress and hunger levels. At the end of the day, users receive a notification asking if they recorded all of their slips for the day and, if not, to record any remaining slips. In a usability study, we evaluated the use and acceptability of Slip Buddy in 16 adults over 4 weeks. Participants used the app 26.8 out of 28 days, tracked about 14 slips, and lost on average 1.5% (SD 0.7%) of their baseline weight, even though no other intervention was provided. In that study, we then used Slip Buddy data to generate predictive models that informed intervention messages to display to the user when they were in a situation in which overeating was triggered in the past. Participants used Slip Buddy for another month while receiving these messages, but participant feedback revealed that they did not find the messages more helpful than slip tracking alone, so we dropped that aspect of the app. As Slip Buddy was not designed as a standalone weight loss intervention, but instead as an alternative to calorie tracking tools used in behavioral weight loss interventions, the next step is to examine the feasibility and acceptability of Slip Buddy as a replacement for the traditional calorie tracking app in a web-based behavioral weight loss intervention (Trials Registration: Clinicaltrials.gov NCT02615171).

## Objectives

This study is a pilot feasibility randomized controlled trial in which participants with overweight or obesity were randomized to receive either the Slip Buddy app or a commercial calorie tracking app during a 12-week counselor-led web-based weight loss intervention. Our first aim is to compare groups on retention and app use to explore whether the commercial app is significantly superior to Slip Buddy, which would point to the need for further modification to Slip Buddy before proceeding to a fully powered efficacy trial. Our second aim is to describe the total number of slips reported and the contextual factors reported at slips, including the location of slips, type of eating episode (eg, lunch and snack), stress, and hunger/satiety. This aim is descriptive in nature. Our third aim is to assess the usability, acceptability, and burden of the Slip Buddy app quantitatively and via qualitative interviews where participants shared what they liked and disliked about the app and the features they would like added. Similar to aim 1, we tested whether the commercial app was significantly superior in terms of usability, acceptability, and burden, which would signal areas for further modification to Slip Buddy before proceeding to an efficacy trial. Our fourth aim is to describe the percentage of weight loss from baseline to 12 weeks in both groups, including the proportion of participants who lost clinically significant weight. This aim is exploratory because only a fully powered efficacy trial could address this question.

## Methods

### Study Design, Settings, and Participants

We conducted a pilot feasibility randomized trial in which participants who were overweight or obese were recruited into a remotely delivered intervention via web-based advertisements at the University of Connecticut, ResearchMatch, and Facebook groups across the United States between July and October 2019. All work was approved by the University of Connecticut Institutional Review Board. We recruited people interested in losing weight with BMI between 27 and 45 kg/m<sup>2</sup>, aged 18-65 years, who had an Android smartphone, and who had phone connectivity at home and work. Exclusion criteria were inability to walk unaided for one-fourth mile without stopping, not being a daily Facebook user (because the group-based part of the intervention was delivered via Facebook), taking medications known to affect appetite and/or weight, having a condition that precludes dietary changes (eg, ulcerative colitis), type 1 or 2 diabetes, gastric bypass surgery or plans to do so during the study period, pregnancy or lactation, severe mental illness or substance use disorder, binge eating disorder, or loss of 5% or more weight in the past 3 months. Recruitment ads contained a link to a screening survey that included a study description, initial informed consent, and screening questions. Participants eligible after the screening survey were emailed the consent form and completed a telephone screening call. The screening call included reviewing the consent form, any remaining eligibility-related questions, and an emailed link to the baseline survey.

Before randomization, potential participants were required to attend an orientation webinar in which study staff used a

methods-motivational interviewing approach to help them understand the scientific rationale of the trial design, research questions, and methods. This helps participants understand the commitment entailed in trial enrollment and helps set clear expectations (eg, transparency about the length of assessments), explain the scientific rationale for procedures (eg, randomization and feasibility versus efficacy testing), diffuse ambivalence about research participation using motivational interviewing techniques, and make explicit commitments to self and trial methods. Upon completion, those interested in proceeding with the study were mailed a Wi-Fi scale (Fitbit Aria) and asked to provide staff with log-in information for the scale so that weight could be recorded for assessments. After randomization, participants had a 60-minute call with a study staff person to receive guidance on how to download and use their assigned app and enter their assigned Facebook group. Participants were allowed to keep the scale and were compensated for completing the assessments.

### Intervention Conditions

Participants in both conditions were assigned one of the diet tracking apps and received the Diabetes Prevention Program (DPP) lifestyle intervention delivered within a counselor-led private Facebook group that included all participants in their respective conditions. Participants randomized to the Slip Buddy condition were provided the Slip Buddy app, and participants randomized to the calorie tracking condition were instructed to install the free, commercially available MyFitnessPal app. Each group had a different counselor who was either a registered dietitian or a clinical psychologist, and each was trained in the app assigned to their respective conditions and led the Facebook group for that condition.

As in our previous work [17,18], the lifestyle intervention was delivered via twice daily posts, and each week's content was based on the corresponding module of the DPP. The DPP assigns participants the goals of (1) calorie tracking to achieve a calorie goal based on the amount needed to lose 1-2 lb (0.45-0.90 kg) per week (modified to weekly slip tracking in the Slip Buddy condition), (2) developing a healthy diet consistent with the American Heart Association guidelines, (3) engaging in 150 to 300 minutes per week of moderate-intensity exercise, (4) developing a strength training regimen consistent with the National Guidelines for Physical Activity, and (5) losing 1-2 lb per week. All DPP content related to dietary self-monitoring was modified in the Slip Buddy condition to focus on slip tracking (as opposed to calorie tracking). Goal setting occurred on Monday mornings when the counselor asked participants to set 2 to 3 diet and exercise goals and gave specific suggestions based on the topic of the week (eg, self-monitoring, reducing added sugar, and adding 15 minutes of exercise) to help participants progress toward their weight loss goal. On Fridays, the counselor put up a weigh-in post asking participants to reply with their weight change in pounds (eg, lost 1 pound) for the week. This ensures participants weigh themselves at least once a week on the study scale and allows an opportunity for problem solving for those who do not lose weight. This procedure is in place of private weigh-ins that typically occur between the counselor and each participant before group meetings in clinic-based programs. Conducting the weigh-ins as a discussion

thread requires less time than having the counselor do individual weigh-ins for each participant, which is a procedure that does not scale well in groups that are as large as 32 participants. Goal accountability occurred every Sunday when the counselor asked participants to report how they did on their weekly goals. In between these key posts were posts related to the topic of the week (eg, nutrition and making time for exercise). On the remaining days, the counselor posted discussion threads relating to that week's module. Each week, staff produced weight and engagement reports for the counselors so they could identify participants who had not engaged in the past week and/or who were not losing weight, and they attempted to engage them in the group by tagging them in posts. Tagging a participant in a post results in them receiving a notification on their Facebook account that, when clicked, leads them to the post in which they are tagged. In our previous studies, this Facebook-delivered weight loss intervention produced mean weight losses at 12 weeks ranging from 2.6% to 4.8% [18,19].

### **Slip Buddy App**

As described above, we developed the Slip Buddy app, which assists users in tracking nonhomeostatic eating and the contextual factors surrounding it (Figure 1) [16]. As *nonhomeostatic eating* is scientific jargon, the app refers to these episodes as *slips*. Participants were instructed to hit an *Oops!* button each time they had a diet slip, defined as any eating that resulted in consuming (food or drink) more than planned at a meal or between meals, eating in the absence of hunger (eg, ate a donut someone brought to work), emotional eating, eating past the point of fullness, or an unhealthy food choice (eg, stopped for fast food instead of cooking). The definition of a slip appeared near the *Oops!* button as a reminder to the user. To increase awareness of when slips are most likely to occur, the app passively collected the date and time for each slip reported. For each slip, participants were asked to rate their

stress and degree of hunger to satiety on a 0 to 10 scale to describe the context of the slip from drop-down menus, including the type of eating episode (eg, lunch and snack) and activity during the episode (eg, working, socializing, and watching television). They were also asked to type in their location (eg, restaurant), food consumed, and any other notable details they wanted to remember later in open text boxes. A check-in tab asks participants to report weight, hours slept last night, stress and hunger in the morning, and stress and hunger in the afternoon, but unlike our first pilot study, we removed the notifications for the morning and afternoon check-ins to keep notifications to a minimum. The only notification occurred at the end of the day, asking participants if they missed entering any slips for the day and, if so, to record the missed slips. The data collected by the app were securely sent to the remote Slip Buddy database server in addition to being recorded in the local database on participants' mobile phones. The history tab showed participants' past slip entries so they could look for patterns in contextual factors such as stress ratings, hunger level, activities, and/or location (Figure 1). Just as the calorie tracking group was given guidance on how to learn from calorie tracking, each week, the counselor instructed Slip Buddy participants to view their slip history from the previous week and use that information to set goals around how to avoid and/or manage cues associated with past slips. For example, if most slips occurred while watching television in the evening, they could set the goals of planning healthy snacks at this time or reducing television time. Participants were urged to use the app to learn when and why they slip and to reduce their slips over time toward the goal of losing 1-2 lb per week. In the Facebook group, the DPP content related to calorie tracking and cues was modified to address slip tracking and to draw participants' attention to the eating cues they were learning about from their Slip Buddy history. No additional content or other modifications were made to the DPP content in the Slip Buddy condition.



Figure 1. Slip Buddy Screen Shots.



**Calorie Tracking Condition**

Participants randomized to the calorie tracking condition were instructed to download MyFitnessPal, a free, commercially available mobile app that provides users with a personalized calorie goal and allows them to track their caloric intake and energy expenditure via exercise in an effort to stay within that goal. Participants were asked to enter everything they eat and drink throughout each day and all of their structured physical activity. They were asked to stay within their calorie goals to facilitate a weight loss of 1-2 lb per week. The group counselor instructed participants to use MyFitnessPal daily and to inspect

their dietary entries for high-calorie foods that could be eliminated to achieve their calorie goal.

**Measures**

**Retention**

Retention was defined as the percentage of participants in each condition who completed the 12-week follow-up measures, which included weigh-in and a survey.

**App Use**

We report 3 metrics of app use: (1) whether participants used their assigned app at least once during the 12-week intervention,

(2) the total number and percentage of possible days participants used their app over the 12 weeks, and (3) whether participants used their assigned app at least once in week 12 (ie, sustained engagement). As the nature of diet tracking differed between the 2 treatment conditions, how we assessed app use also differed. For participants in the Slip Buddy condition, we intended to categorize participants as having used the Slip Buddy app for a given day if (1) backend data from the app revealed at least one slip was recorded or check-in completed (optional) or (2) in the absence of slips, the participants responded to the end-of-day check-in saying that they did not have any slips that day. Staff reviewed the data in the Slip Buddy database server (ie, backend data) and recorded the number of days each week each participant used the app (eg, either recorded a slip or responded to a notification or check-in reporting that they experienced no slips). However, some participants reported that they did not see or receive the end-of-day check-in notifications from the app, which we determined was related to the authentication token on the phone expiring periodically. The end-of-the-day notification gives the participant the opportunity to confirm that no slips occurred if none had been recorded thus far. Without the notification, we could not distinguish between a day in which the participant did not track slips and a day in which no slips occurred. For this reason, backend data would be an underestimation of app use. As failure can sometimes occur while transmitting app data to the remote database server, we also collected self-report app use data by emailing participants each week a single item asking them how many days they used the app to track slips that week. Self-report data were available for 74.4% (282/379) of weeks across all participants (counting only 7 weeks for the participant who withdrew because of pregnancy). As 26% of self-reported data were missing and backend data were incomplete by an unknown amount, we leveraged both forms of data to measure app use. We used the larger of the 2 values for 2 reasons: (1) when self-reported data are higher than backend data, it could correct for the underestimation bias of backend data and (2) when backend data are higher than self-report, it could correct for recall bias from self-report. The weakness is that we do not have a way to correct for recall bias from self-report that overestimates use, which surely exists to some extent. Self-report data were used for 58.1% (220/379 weeks) of the total weeks, and backend data were used for 27.4% (104/379 weeks) of weeks, which includes the 7 weeks when the backend data were higher than the self-report data and the 97 weeks in which self-report data were missing. On the remaining 14.5% (55/379 weeks) of weeks, self-report and backend data were the same so that the value was used. Although there is no way to correct for possible overestimations via self-report, on 25.6% (97/379 weeks) of weeks, only backend data were available, which would be an underestimate for those weeks. As a sensitivity analysis, we calculated app use using self-reported data if available, and when self-report data were not available, backend data were used. These metrics only differ from the main analysis for the 7 weeks (1 week for each of the 7 participants) where app use abstracted from the backend was higher than self-reported app use.

For participants in the calorie tracking condition, research staff reviewed MyFitnessPal records, and we coded a *complete day*

of calorie tracking any day in which participants tracked 2+ meals and 800+ kcal/day, as has been done elsewhere [20,21]. As participants in the calorie tracking condition were instructed to track all food and beverage intake, we only included *complete days* of tracking in our calculations of MyFitnessPal use.

Using the above definitions, we calculated the number and percentage of days participants in each treatment condition used their assigned app over the 12-week intervention. As the Slip Buddy app was down for 2 days in week 3, participants in this condition could have only used the app on a maximum of 82 days versus the 84 possible days for participants in the calorie tracking condition. We also categorized participants in both conditions as to whether they used their assigned app at least once over the 12-week intervention and whether they used their assigned app during week 12 as a measure of sustained engagement. Two participants were withdrawn or dropped out of the intervention because of incident pregnancies. For these women, app use was not assessed after they were no longer in the intervention (after week 7 for the Slip Buddy participant who became pregnant and after week 2 for the calorie tracking participant); instead, the calculation of percentage of days the app was used only counted days they were in the intervention.

### **Slip Buddy App Data (Slip Buddy Participants Only)**

#### ***Slips Reported***

Backend data from the app were used to describe the number of slips reported for each participant during the intervention period and contextual factors related to slips.

#### ***Location of Slip***

Participants were asked to enter a note about where they were when the slip occurred. The first author collapsed responses into categories that included work, home, other person's house, restaurant/bar, at an event (eg, football game), in the car, or at the gym.

#### ***Nature of Eating Episode***

Participants also indicated the nature of the eating episode in which the slip occurred, which included the choices of breakfast, lunch, dinner, dessert, snack, or alcohol. Alcohol was included to capture drinking episodes that happen outside of the context of meals or snacks and to prompt participants to think of excess alcohol intake as a dietary slip.

#### ***Activity During Slip***

Participants also indicated what they were doing from a drop-down menu of domestic activities (eg, chores), working/studying (eg, employment and school), socializing, screen time, or commuting.

#### ***Stress and Hunger or Fullness Ratings of Slips***

When they entered a slip, participants rated how much stress they were experiencing before their slip on a scale of 0 to 10, where 0 indicates no stress and 10 indicates extreme stress. Stress scores of 5 and above were considered moderate to high stress, whereas stress ratings of less than 5 were considered low stress. Participants also rated how hungry or full they felt before they slipped on a scale of 0 to 10, where 0 indicates extremely

hungry, 5 indicates comfortably full, and 10 indicates stuffed, that is, uncomfortably full.

### **MyFitnessPal Data (Calorie Tracking Participants Only)**

The participants were asked to record their diet every day for 12 weeks. These data were extracted from MyFitnessPal and coded for analysis. The first level of coding included recording each day that the participant entered at least one item. The second level extracted the number of eating episodes and calories each day.

### **Usability**

The System Usability Scale (SUS) [22] was used at 12-weeks to assess the Slip Buddy app's usability. The SUS is a 10-item 5-point Likert scale questionnaire regarding human-computer interaction. For a participant who only answered 9 of the 10 questions, we used their mean of those 9 items to impute a response to the tenth item. An SUS score above 70 is considered acceptable and above average, whereas a score above 85 is considered excellent [23]. Moreover, when users rate a system with an SUS score of 82 (SD 5), they tend to be *promoters* of the system, which means they are likely to recommend it to a friend [24].

### **Acceptability**

At 12 weeks, participants in both conditions rated the helpfulness and ease of use of their assigned app (response options: strongly disagree, disagree, neutral, agree, or strongly agree). We dichotomized responses as strongly agree/agree versus strongly disagree/disagree/neutral. As the Slip Buddy app is exclusively focused on diet, unlike calorie tracking apps that address both diet and exercise, we included a follow-up question asking participants to rate whether a feature that would allow them to track exercise slips (ie, times when they had planned to exercise but did not follow through) would increase the effectiveness of Slip Buddy app. Acceptability was also evaluated in postintervention focus groups via 2 questions: "what did you like most about Slip Buddy app and why?" and "what did you like least about Slip Buddy app and why?" During the intervention, participants started a discussion about the possibility of the Slip Buddy app having a feature that would allow people to track when they were tempted to slip but resisted that temptation. Given the enthusiasm for the idea, we added a question to the focus group script asking participants about the extent to which they would like to track temptations that did not turn into slips.

### **Burden**

At 12 weeks, participants in both conditions rated how burdensome it was to use their assigned app on a scale of 0 to 100, with 0 being not at all burdensome and 100 being very burdensome. Participants rated how much they agreed that the app was time consuming, taxing, and tedious (response options: strongly disagree, disagree, neutral, agree, or strongly agree). We dichotomized responses as strongly agree/agree versus strongly disagree/disagree/neutral.

### **Weight**

Weight was obtained at baseline and at 12 weeks via the Wi-Fi scales sent to participants upon enrollment.

### **Participation Engagement in the Facebook Group**

Participant engagement is defined as participant posts, replies, reactions (eg, love, wow, angry, and sad), and participation in intervention polls, which are used either as a way of assessing participant knowledge (eg, pop quizzes) or as a way for participants to share their diet and/or exercise barriers. We extracted engagement data from the private Facebook group using the Grytics app, except poll data, which were manually extracted because Grytics does not capture poll data. We summarized the total number of original posts, replies, reactions, and polls that each participant participated in. In addition, we calculated the percentage of participants in each condition who replied to each of the 12 weekly weigh-in posts.

### **Statistical Analysis**

We summarized retention, app use, slips, usability, acceptability, burden, and engagement in the Facebook groups, including the percentage participating in the weekly weigh-ins using descriptive statistics. For variables that were normally distributed, we described distributions using mean and SD, and for variables that were not normally distributed, we described distributions using median and IQR. We compared use, retention, usability, acceptability, and engagement by treatment condition using *t* tests, chi-square tests, Fisher exact tests, or Mann-Whitney *U* test as appropriate. We compared the treatment conditions on app burden using the Wilcoxon rank-sum test. Insufficient retention, acceptability, and use were assumed if the calorie tracking condition showed a statistically significant advantage relative to Slip Buddy. Statistical tests were not used to compare groups on weight loss because this pilot study was not powered for weight loss efficacy; thus, it is not appropriate to perform such tests, as discussed elsewhere [25]. We used an intent-to-treat approach to describe the weight change. Two participants (1 in each condition) became pregnant during the study period. We used the latest available prepregnancy weights (from weeks 2 and 3 for the 2 participants, respectively) from their study scales as their follow-up values. Three participants did not provide weight during the follow-up. We also used their latest weight from their study scales (weeks 6, 9, and 10, for the 3 participants, respectively) as follow-up values. We secondarily reported weight loss assuming no weight loss for the 2 participants who became pregnant and the 3 participants lost to follow-up (ie, baseline observation carried forward), and we secondarily reported weight change excluding these 5 participants for whom we did not have nonpregnant follow-up weight. We conducted a directed content analysis [26] of the focus group data on acceptability. The first author developed a codebook based on themes emerging from the participant responses. Two coders independently coded responses, and discussion was used to achieve consensus on disagreements. Interrater reliability (IRR) was also calculated [27]. We summarized the frequency of the themes. Data management and quantitative analyses were conducted using SAS 9.4 (SAS Institute Inc).

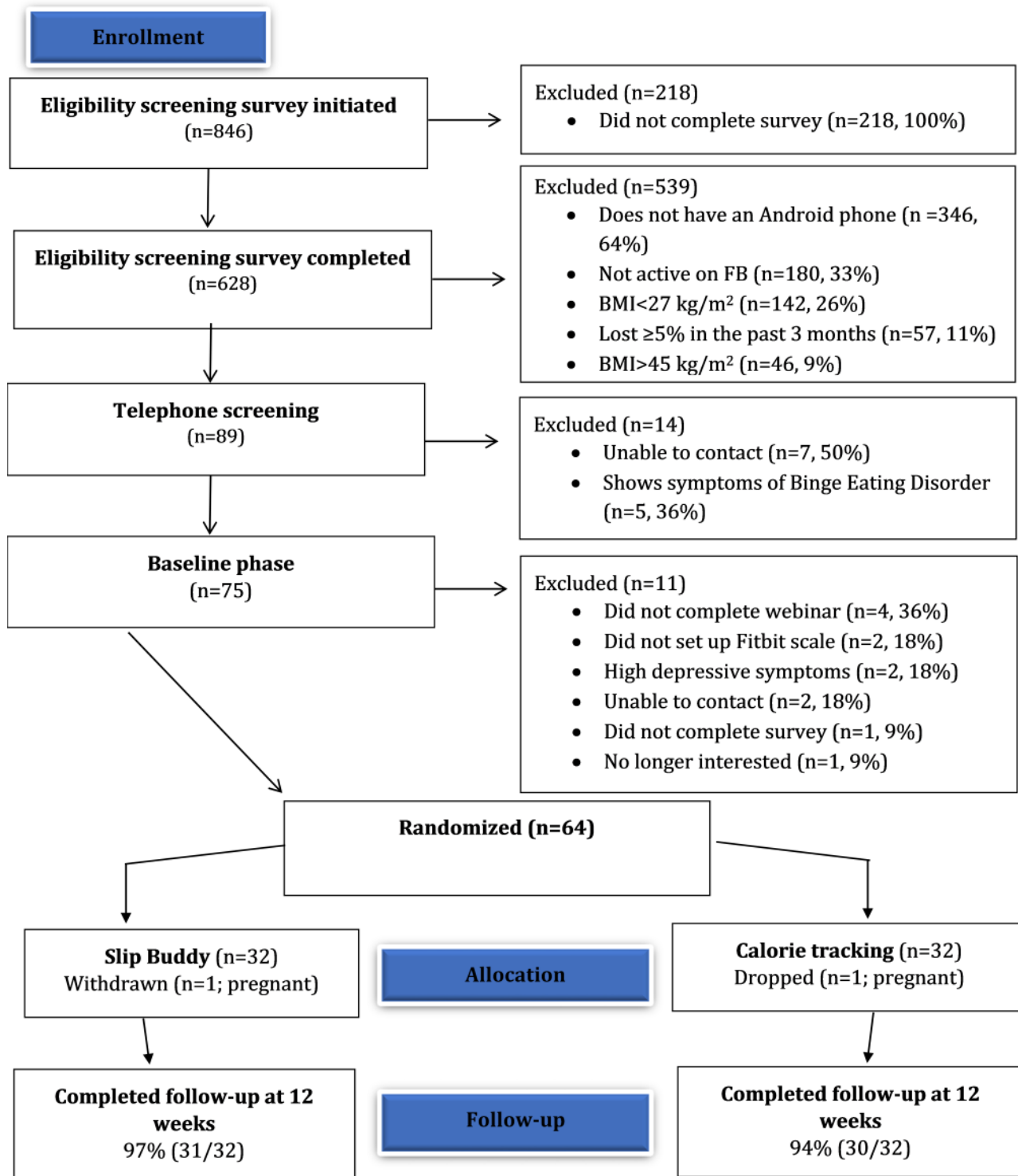
## Results

### Recruitment

A total of 846 individuals initiated the eligibility screening survey (Figure 2). Among individuals screened for eligibility, the most common reasons for exclusion were not owning an Android phone, not being an active Facebook user, BMI outside

the eligible range, or recent weight losses of 5% or more (Figure 2). We randomized the 64 participants to 1 of 2 treatment conditions. Overall, participants were, on average, 39.8 (SD 11.0) years old, with a baseline BMI of 34.2 (SD 4.9) kg/m<sup>2</sup>; 75% (48/64) were female; and 81% (52/64) were non-Hispanic White. Participants lived in 18 US states, and 56% (36/64) of the participants were from Connecticut. Additional characteristics are presented in Table 1.

Figure 2. Consolidated Standards of Reporting Trials diagram. FB: Facebook.



**Table 1.** Characteristics of participants by treatment condition (N=64).

Characteristics	Treatment condition	
	Slip Buddy (n=32)	Calorie tracking (n=32)
Age (years), mean (SD)	39.5 (9.6)	40.2 (12.3)
BMI at enrollment (kg/m <sup>2</sup> ), mean (SD)	34.9 (5.3)	33.4 (4.4)
Female, n (%)	24 (75)	24 (75)
<b>Race, n (%)</b>		
Non-Hispanic White	26 (81)	26 (81)
Hispanic or Latino (any race)	1 (3)	4 (13)
Non-Hispanic Black	3 (9)	1 (3)
Non-Hispanic Asian	N/A <sup>a</sup>	1 (3)
Other race or multiracial	2 (6)	N/A
<b>Marital status, n (%)</b>		
Married or living with a partner	24 (75)	22 (69)
Single	8 (25)	7 (22)
Divorced or separated	N/A	3 (9)
<b>Education, n (%)</b>		
At most high school	2 (6)	N/A
Trade or technical school, some college, or associate degree	8 (25)	10 (31)
Bachelor's degree	9 (28)	4 (13)
Some graduate coursework	4 (13)	7 (22)
Graduate degree	9 (28)	11 (34)
<b>Employment status, n (%)<sup>b</sup></b>		
Employed full time	22 (69)	25 (78)
Employed part time	8 (25)	5 (16)
Homemaker (not looking for a job)	2 (6)	N/A
Student	1 (3)	4 (13)
Retired	N/A	1 (3)

<sup>a</sup>N/A: not applicable.

<sup>b</sup>Participants could select more than 1 employment status. In the Slip Buddy condition, 1 participant was employed part time and a student. In the calorie tracking condition, n=1 was employed part time and a student, n=1 was employed full time and a student, and n=1 was retired and employed part time.

## Retention

Retention was high in both treatment conditions, with 97% (31/32) of Slip Buddy participants and 94% (30/32) of calorie tracking participants providing follow-up data ( $P>.99$ , Fisher exact test).

## App Use

Nearly all participants randomized to the Slip Buddy condition (31/32, 97%) and the calorie tracking condition (31/32, 97%) used their assigned app at least once during the 12-week intervention. Participants in the Slip Buddy condition used their assigned app on a mean of 44.0 (SD 25.8) days. Participants in the calorie tracking condition used their app on a mean of 46.3 days (SD 24.0), which represented use on an average of 53.8% (SD 31.3%) of possible days for participants in the Slip Buddy condition and 57.5% (SD 28.4%) of possible days for

participants in the calorie tracking condition ( $t_{62}=0.495$ ;  $P=.44$ ). In terms of sustained use of their assigned app, 55% (18/31) of Slip Buddy participants used the app at least once in week 12 of the intervention compared with only 35% (11/31) of calorie tracking participants ( $X^2_1=2.4$ ;  $P=.13$ ). The proportion of participants using their assigned app each week of the 12-week intervention period is shown in [Multimedia Appendix 1](#).

App use by Slip Buddy participants was nearly identical in a sensitivity analysis assessing app use by self-report data when available, and when self-report data were not available, backend data were used. Overall, 97% (30/31) of participants used the app at least once during the 12-week intervention. They used the app on an average of 43.8 (SD 25.7) days, representing 53.4% (SE 31.2%) of possible days; 55% (17/31) of the participants used the app at least once in week 12 of the intervention.

## Slips Reported

One participant did not use the Slip Buddy app, and 1 participant responded to the notification but did not record any slips. The remaining participants reported a total of 538 slips during the 12-week intervention period. Participants who reported slips ( $n=30$ ) reported a median of 15 slips (IQR 8-23; range 2-66; mean 17.9, SD 14.4). Most of slips happened at home (297/538, 55.2%), followed by work (113/538, 21.0%) and restaurant/bar (86/538, 16.0%; Table 2). The nature of the eating episode most likely to be reported as a slip was a snack (220/538, 40.9%),

followed by dinner (102/538, 19.0%), dessert (77/538, 14.3%), and lunch (60/538, 11.2%; Table 2). Activities engaged in when the slip occurred were split over work or studying (153/538, 28.4%), socializing (130/538, 24.2%), screen time (123/538, 22.9%), and domestic activities (112/538, 20.8%), and a small percentage of slips occurred while commuting (20/538, 3.7%; Table 2). The median stress rating during slips was 4 (IQR 2-5). The median hunger/fullness rating during slips was 4 (IQR 3-5). One-fifth of the slips (106/538, 20%) occurred when both stress and hunger were low, and another 20% (105/538) of the slips occurred when both stress and hunger were high.

**Table 2.** Location, activity, eating episode, stress, and satiety associated with slips reported by participants over 12 weeks ( $N=538$  slips).

Slip characteristics	Value, n (%)
<b>Location</b>	
Home	297 (55.2)
Work	113 (21)
Restaurant or bar	86 (16.0)
Another person's house	22 (4.1)
The car	10 (1.9)
An event	8 (1.5)
The gym	2 (0.4)
<b>Activity</b>	
Work or studying	153 (28.4)
Socializing	130 (24.2)
Screen time	123 (22.9)
Domestic activities	112 (20.8)
Commuting	20 (3.7)
<b>Eating episode</b>	
Snack	220 (40.9)
Dinner	102 (19.0)
Dessert	77 (14.3)
Lunch	60 (11.2)
Breakfast	40 (7.4)
Alcohol	39 (7.3)
<b>Stress</b>	
Lower range, 0-4	298 (55.4)
Higher range, 5-10	240 (44.6)
<b>Hunger to satiety</b>	
Hungry range, 0-4	299 (55.6)
Full range, 5-10	239 (44.4)

## Usability

The mean SUS score for the Slip Buddy condition was 64.8 (SD 16.5), which is the marginally acceptable range. Comparatively, the mean SUS score for participants in the calorie tracking condition was 76.3 (SD 17.6), which is considered good acceptability. Participants in the calorie tracking condition rated the MyFitnessPal app as more usable

on average than the Slip Buddy participants rated the Slip Buddy app ( $t_{59}=2.64$ ;  $P=.01$ ). Usability issues reported during the intervention included the Slip Buddy app crashing on some phone models and a temporary outage, both of which were fixed during the study.

## Acceptability

Among participants who completed the follow-up survey, 39% (12/31) of Slip Buddy participants agreed or strongly agreed

that tracking slips was helpful, whereas 77% (23/30) of calorie tracking participants agreed or strongly agreed that tracking diet and exercise was helpful ( $\chi^2_{1,N=61}=8.9$ ;  $P=.003$ ; Table 3).

**Table 3.** Acceptability of assigned tracking app by treatment condition.

Acceptability item <sup>a</sup>	Slip Buddy (n=31), n (%)	Calorie tracking (n=30), n (%)	$\chi^2$ (df)	P value
Tracking my slips with Slip Buddy app/tracking my diet and exercise with MyFitnessPal was helpful for me	12 (39)	23 (77)	8.9 (1)	.003
Tracking diet slips on Slip Buddy app/MyFitnessPal is easy	24 (77)	26 (87)	0.9 (1)	.35
Using the Slip Buddy app/MyFitnessPal is tedious	9 (29)	12 (40)	0.8 (1)	.37
Using the Slip Buddy app/MyFitnessPal is taxing	5 (16)	8 (28)	1.2 (1)	.28
Using the Slip Buddy app/MyFitnessPal is time consuming	6 (20)	13 (43)	3.8 (1)	.05

<sup>a</sup>Proportion of participants responding with *strongly agree* or *agree* versus *strongly disagree*, *disagree*, or *neutral*.

Most of both Slip Buddy (24/31, 77%) and calorie tracking participants (26/30, 87%) agreed or strongly agreed that using their respective app was easy ( $\chi^2_1=0.9$ ;  $P=.37$ ; Table 3). Two-thirds of Slip Buddy participants (21/31, 68%) agreed or strongly agreed that adding the ability to track exercise slips would be helpful.

## Burden

On a scale of 0 to 100, the median burden rating for participants in the Slip Buddy condition was 30 (IQR 15-50), and the median burden rating for participants in the calorie tracking condition was 45 (IQR 10-60; Mann-Whitney  $U$  test,  $U=949.5000$ ;  $P=.78$ ). The proportion of participants who agreed or strongly agreed that using their assigned app was tedious or taxing did not differ by treatment condition (Table 3). Finally, 20% (6/31) of Slip Buddy participants agreed or strongly agreed that using the Slip Buddy app was time consuming, whereas 43% (13/30) of MyFitnessPal participants agreed or strongly agreed that using MyFitnessPal was time consuming ( $\chi^2_1=3.8$ ;  $P=.05$ ; Table 2).

## Feedback From Participants in the Slip Buddy Condition

A total of 88% (28/32) of participants in the Slip Buddy condition attended postintervention focus groups, and they made a total of 35 responses to the question about what they liked most about the Slip Buddy app (IRR=94%;  $\kappa=0.89$ ). The most common theme of responses was the ease of use/simple concept (23/28, 66%), followed by increasing accountability and/or awareness of overeating and/or triggers (8/28, 23%), the end-of-day reminder to track slips (2/28, 5%), feeling motivated to not slip so there would be nothing to track (1/28, 3%), and other (1/28, 3%). Participants made a total of 35 responses about what they liked least about the Slip Buddy app (IRR=91%;  $\kappa=0.89$ ). Responses reflected the following themes: technical issues (eg, app crashing and notifications not going away; 10/35, 29%), easy to forget to use or not sure how to use when no slips (7/35, 21%), did not find relevant stress ratings (5/35, 15%), focus on slips was too negative (4/35, 12%), was not sure what

to count as a slip (4/35, 12%), did not include diet instruction (4/35, 12%), and slip history screen was not as informative as it could be (1/35, 3%). For the final question regarding their thoughts on a feature that would allow them to track when they were tempted but did not slip, 75% (21/28) said they would be enthusiastic about this feature. The remainder said they worried that it would add too much burden.

## Weight Change

Over 12 weeks, participants randomized to the Slip Buddy condition had an average weight loss of  $-6.5$  lb (SD 9.7) or 3.0% (SD 4.5%) of their baseline weight, and participants randomized to the calorie tracking condition had a weight loss of  $-7.5$  lb (SD 10.7) or 3.6% (SD 4.9%) of their baseline weight (Table 4). In terms of clinically significant weight loss, 31% (10/32) and 34% (11/32) of participants randomized to the Slip Buddy and calorie tracking conditions, respectively, achieved 5% or greater weight loss, and 47% (15/32) and 47% (15/32), respectively, achieved 3% or greater weight loss (Table 4). In a secondary analysis assuming no weight loss for the 2 participants who became pregnant and the 3 participants lost to follow-up (ie, baseline observation carried forward approach), weight losses were on average  $-6.1$  (SD 9.8) lb and  $-2.8\%$  (SD 4.6%) among Slip Buddy participants and on average  $-7.1$  (SD 10.8) lb and  $-3.4\%$  (SD 5.0%) among calorie tracking participants, with 31% (10/32) and 34% (11/32) of participants, respectively, losing 5% or more weight compared with baseline, and 44% (14/32) and 44% (14/32), respectively, losing 3% or more weight compared with baseline. In a secondary analysis of the 59 participants who provided weight at follow-up and were not pregnant ( $n=30$  Slip Buddy and  $n=29$  calorie tracking), weight losses were on average  $-6.5$  (SD 10.0) lb and  $-3.0\%$  (SD 4.7%) among Slip Buddy participants and on average  $-7.9$  (SD 11.1) lb and  $-3.8\%$  (SD 5.1%) among calorie tracking participants, with 33% (10/30) and 38% (11/29) of participants, respectively, achieving 5% or greater weight loss from baseline and 47% (14/30) and 48% (14/29), respectively, achieving 3% or greater weight loss.

**Table 4.** Weight change from baseline to 12 weeks, by treatment condition.

Weight variables <sup>a</sup>	Slip Buddy (n=32)	Calorie tracking (n=32)
Baseline weight (lb) <sup>b</sup> , mean (SD)	217.2 (39.0)	208.5 (37.2)
Follow-up weight (lb), mean (SD)	210.7 (39.5)	201.0 (36.3)
Absolute weight change (lb), mean (SD)	-6.5 (9.7)	-7.5 (10.7)
Percentage weight change, mean (SD)	-3.0 (4.5)	-3.6 (4.9)
5% or greater weight loss, n (%)	10 (31)	11 (34)
3% or greater weight loss, n (%)	15 (47)	15 (47)

<sup>a</sup>We used the last available weight from the study scales for 8% (5/64) participants. In Slip Buddy condition, 3% (1/32) participant became pregnant (used week 3 weight), and 3% (1/32) participant did not provide follow-up weight (week 6 weight). In the calorie tracking condition, 3% (1/32) participant became pregnant (we used week 2 weight), and 6% (2/32) participants did not provide follow-up weight (we used weeks 9 and 10 weight).

<sup>b</sup>1 lb = 0.45 kg.

### Participant Engagement in the Facebook Group

In the Slip Buddy condition, the median total replies per participant was 55.00 (IQR 11.75-79.00), which was not significantly different from the median total replies of 29.5 (IQR 17.25-60.75;  $U=445.5000$ ;  $P=.37$ ) in the calorie tracking condition. In the Slip Buddy condition, participants reacted to a median of 13.00 (IQR 3.25-47.75) posts or replies, which was not significantly different from the median reactions of 13.00 (IQR 4.00-18.75;  $U=462.500$ ;  $P=.51$ ) in the calorie tracking condition. Few participants posted original posts (31% (10/32) in Slip Buddy and 34% (11/32) in calorie tracking); the median number of original posts participants made was 0 (IQR 0-1) in both conditions ( $U=508.000$ ;  $P=.95$ ). Finally, the Slip Buddy condition participants had a median total poll votes of 12.00 (IQR 6.25-16.00) compared with 8 (IQR 6.25-13.75) in the calorie tracking condition, a difference that was not statistically significant ( $U=440.500$ ;  $P=.33$ ).

In terms of weekly weigh-in participation, on average, 58.07% (SD 14.32; range 34%-81%) of Slip Buddy participants and 51.30% (SD 18.48; range 25%-88%) of calorie tracking participants replied to weigh-in posts each week ( $t_{1,22}=1.003$ ;  $P=.33$ ). Week 1 had the highest participation in both groups (26/32, 81% and 28/32, 88% in Slip Buddy and calorie tracking, respectively), and participation declined over time to 63% (20/32) and 56% (18/32), respectively, by week 6 and to 34% (11/32) and 25% (8/32), respectively, by week 12.

## Discussion

### Principal Findings

Findings revealed that although participants in both treatment conditions used their assigned apps on a similar percentage of intervention days (ie, 54% of days for participants in the Slip Buddy condition and 58% of days for participants in the calorie tracking condition), 55% (17/31) of Slip Buddy participants used the app at week 12 of the intervention compared with only 35% (11/31) of calorie tracking participants. However, this difference was not statistically significant. Less than one-third of Slip Buddy participants agreed that using the Slip Buddy app was tedious (9/31, 29%), taxing (5/31, 16%), or time consuming (6/31, 20%); 77% (24/31) agreed that tracking slips was easy;

but only 39% (12/31) agreed that tracking their slips with the app was helpful, which was significantly lower than that in the calorie tracking app condition. Slip Buddy also received lower usability ratings than the commercial calorie tracking app, perhaps not surprisingly, as commercial apps are years ahead of Slip Buddy in user experience optimization. Slip Buddy participants reported barriers such as technical difficulties, forgetting to use the app if they had not slipped in a while, and finding the exclusive focus on slips to be too negative. Slip data revealed that most of the slips reported happened at home, followed by work, and that snacks and dinner time were the eating episodes at which slips were most likely to occur. Activities that co-occurred with slips were distributed fairly evenly across work, socializing, screen time, and domestic tasks. Less than half of the slips occurred under conditions of moderate to high stress and over half occurred while hungry.

The Slip Buddy app was designed to reduce dietary self-monitoring to possibly its simplest form by only necessitating the recording of aberrant eating episodes. Despite its simplicity, use rates in this study appeared fairly comparable with app use in the commercial calorie tracking app condition. Interestingly, a randomized trial that compared a commercial calorie tracking app (Calorie Counter by Fat Secret) with the lower-burden Meal Logger app, which allows users to track their diet by taking photos of what they eat [28], found that the calorie tracking app was used on more intervention days over 6 months than the less intensive photo app (41% vs 28% of intervention days). Research on the adoption of digital health innovations suggests that usefulness and ease of use are 2 major drivers of use; thus, improving the use rates of Slip Buddy may involve enhancing these factors [29]. Accordingly, we are currently using our qualitative findings to guide updates to the app that would improve its usefulness and ease of use. Positive subjective social norms, meaning the belief that other people are using and benefiting from the app, have also been found to affect the use of digital health innovations [30]. People wanting to lose weight are more likely to have been exposed to commercial calorie tracking apps as these apps have been on the market for years and have millions of users, which has not only created social norms around these apps but has also allowed much more time for the user experience to be optimized. Given the long-term dominance of calorie tracking apps in the commercial weight loss space, these apps may also shape user



expectations about what a weight loss app should be, and this could influence how they feel about apps that do not include the features they have come to expect. Indeed, some participants commented in the Facebook group that they would love to have the ability to track both slips and calories, whereas others were glad not to be tracking calories. At follow-up, we asked Slip Buddy participants if they had used a calorie tracking app in the past, and 68% (21/31) of participants said they had. Ultimately, people seem to want choices and flexibility in their options. Given that fatigue can set in with any long-term self-monitoring strategy, the ability to change self-monitoring strategies over time might be optimal.

Our focus group data revealed that the infrequency in which the Slip Buddy app needed to be used (ie, only when a slip occurs) may have led some people to forget to use it. Slip Buddy participants, on average, only recorded an average of 1 to 2 lapses per week, which is far less than daily and possibly insufficient to capture enough calories to lose weight even if all slips were successfully eliminated; however, without calorie data, this is unknown. Studies using ecological momentary assessment to have people track dietary lapses have reported means ranging from 2.7 to 11.8 lapses per week [31]. A recent trial that combined a web-based weight loss program with lapse tracking reported on average 29.7 lapses per participant over 10 weeks for a rate of 2.9 lapses per week over 10 weeks [15]. In that trial, participants were following the WW plan, and lapses were defined as any eating that went over the individual's point target for a meal or snack. Having slips tied to calorie and/or point goals may give users more guidance on how to identify lapses. Our goal was to help people identify lapses in the absence of traditional forms of dietary tracking, so they may need more specific guidance on how to identify lapses. Focus groups revealed that a few participants were not sure what to count as a lapse, which may suggest that our definition of a slip was too narrow or too narrowly interpreted to capture enough eating episodes that could be considered dietary lapses. Alternatively, participants might not count slips that they felt were somehow justified, such as overeating at dinner on a day when a meal was skipped or having an extra slice of cake on one's birthday. Research is needed to determine users' lapse tracking accuracy by comparing lapse tracking data with 24-hour dietary recalls. Such data would also reveal the type of eating episodes people perceive as lapses and inform instructions they are given for how to track dietary lapses. With a rate of only 1 to 2 lapses recorded each week, instructions for lapse recording could be modified to help people capture more of their eating episodes that could be made healthier, not only to facilitate more consistent use but also to increase weight loss. The terms *slip* and *lapse* may be too limiting. Instead, users could be guided to track any eating episode they think has room for improvement or be given a minimum limit of eating episodes to record and work on each week, regardless of whether they consider those episodes to be slips or not. Another possible explanation for our finding of fewer lapses recorded relative to Forman et al [15] is that in their study, participants received 6 notifications per day to remind them to enter data relating to lapse triggers. Additional reminders to track lapses may increase users' awareness of their dietary lapses.

Most participants wanted the Slip Buddy app to include the ability to track instances when they resisted temptations to give them an opportunity to see improvement in their ability to deal with contextual factors that cue slips. Including temptation tracking might allow users to develop a more regular habit of monitoring their eating habits and the circumstances in which they make healthy and unhealthy choices. In another study, participants in a behavioral weight loss program were asked to record their temptations and lapses in a paper diary in the final week of the program. They found that temptations were more likely than lapses to be followed by coping behaviors, suggesting the value of having people record both [12]. By starting with a very simple app and using a user-centered approach, we can now add new features suggested by participants only to the point at which those features continue to add value without undue burden. Most participants were also in favor of adding a feature that would allow them to track when they had exercise slips (ie, skipped a planned workout). Our next iteration will allow users to track these things. As a minority of our participants felt that additional features could undermine the app's simplicity, a user-centered process that is mindful of individual differences in perceived user burden will be important when adding new features to this and any app.

Despite the vast literature on emotional eating [32], only 44.6% (240/538) of slips occurred under conditions of moderate to high stress, and 14% (4/28) of participants in the focus groups said they did not find the stress ratings relevant to them because they do not feel they are a stress eater. Stress is one of many circumstances that can cue nonhomeostatic eating, and in this sample, it did not appear to drive most of the slips. Interestingly, 19.5% (105/538) of slips recorded were under conditions of both low stress and low hunger. The OnTrack app study included a wider range of triggers for users to record, including tiredness, temptations, missed meals/snacks, socializing, television, negative interpersonal interaction, cognitive load, food cues, alcohol consumption, unhealthy food availability, and planning food intake in their slate of possible triggers [15]. To satisfy individual differences and minimize user burden, apps could allow users to customize the cues they want to track by giving them choices from a wide range of options, including both emotional and physical states (eg, boredom and pain). An alternative explanation for a few slips recorded under conditions of high stress could be that stress may cause people to forget about a diet lapse or to be less aware of a diet lapse. Further research should explore how people decide an eating episode is a lapse and whether that changes under different emotional or physiological circumstances. For example, a person who eats a large amount in response to being extremely hungry may not perceive this eating as a lapse because they were very hungry and felt eating that volume of food was necessary under the circumstances. Some studies have users record their emotional and physical states throughout the day, which allows for even higher precision insights into the relationships between these states and dietary lapses [15,33].

This pilot feasibility trial was not powered to detect group differences in weight loss, the primary outcome planned for the larger fully powered randomized trial. Statistical comparisons of clinical outcomes in pilot feasibility trials are also

inappropriate because of the inflated risk of type 1 and type 2 errors [25]. However, weight losses can be considered in the context of other trials of technology-based weight loss interventions. Participants randomized to the Slip Buddy condition lost an average of 6.5 lb in 12 weeks with a retention rate of 97%, which is comparable with the Meal Logger trial discussed above, where participants receiving the photo app lost 4.8 lb at 6 weeks and 5.5 lb at 6 months with a 96% retention rate [28]. In a pilot trial of a podcast and social media-delivered intervention, participants lost 6.4 lb in 12 weeks with a retention rate of 85% [34]. In our pilot trial, 31% (10/32) and 34% (11/32) of participants in the Slip Buddy and calorie tracking conditions, respectively, lost clinically significant weight ( $\geq 5\%$ ). This is comparable with findings from a randomized trial (NCT01479062) of a completely automated weight loss intervention based on the DPP in which 35% of participants lost clinically significant weight at 6 months [35], although our trial was only 12 weeks long compared with 6 months in that study. The mean percentage of weight loss in the Slip Buddy condition (ie, 3%) also fell into the range observed in the OnTrack study (ie, 2.91%-4.65% depending on the version of WW used in combination with OnTrack), which had participants track WW points, dietary lapses, and 16 lapse triggers 6 times a day [15]. Our findings suggest that some people in the Slip Buddy condition successfully lost clinically significant weight, although they were tracking slips and not total calories. Dietary mobile apps are not likely a one-size-fits-all. Further research should explore individual differences that predict who will be successful using different approaches to dietary self-monitoring.

This study had several limitations. First, the sample size was not large enough to compare the 2 conditions for weight loss. However, the purpose of this work was to evaluate the feasibility and acceptability of the Slip Buddy app using a user-informed process to guide improvements to the technology before conducting a fully powered randomized trial. We chose a 12-week intervention length to allow us to gain user insights after a prolonged period of use, but this study does not provide information on tracking habits over more extended periods, such as whether slip tracking decreases over time as participants learn their triggers and slip less. Another limitation is that our

ability to measure app use via backend data was hampered by the fact that some participants did not see or receive end-of-day notifications that, when clicked, would indicate whether no slip entries meant a slip-free day or nonuse. For this reason, we had to rely on self-report use data that are prone to biases because of forgetting or social desirability. This can be ameliorated in the next version, and the addition of temptation and exercise slip tracking will give the user more to do each day with the app, even in the absence of slips. Another limitation is that the Slip Buddy app was not operating for 2 days, which could have affected use in subsequent days to the extent that participants were frustrated by this. As it is a newly developed app, bugs and crashes are more common than commercial apps that have been around for many years, and this will certainly impact usability ratings. An additional limitation is that both Slip Buddy and MyFitnessPal allow users to track other behaviors that may impact weight (eg, exercise, sleep, and mood), and we know little about the use of these features and whether their use impacted outcomes. Finally, a limitation is that the sample overrepresented non-Hispanic White women (40/64, 63% of our sample). In future research, recruitment will need to limit enrollment of non-Hispanic White women, the population segment that too often comprises most of the weight loss trial samples [36], to no more than the proportion of the population they represent.

Although human counseling is associated with better outcomes in technology-delivered behavioral weight loss interventions [37], it is also the main expense and primary barrier to scalability; thus, digital health tools that execute behavioral strategies are needed to reduce the time spent by human counselors. The Slip Buddy app was designed not only to simplify dietary self-monitoring but also to help users identify and disrupt cue-behavior linkages, a process that is typically facilitated via counseling. Data from this study will guide the next iteration of the Slip Buddy app, which will balance the need for simplicity with the need for additional technology-delivered behavioral strategies beyond self-monitoring and feedback. These findings can also inform future studies on how technology can be leveraged to execute simpler forms of dietary self-monitoring, given the high burden associated with traditional calorie tracking apps.

---

## Acknowledgments

This work was funded by National Institutes of Health grants R01HL122302 and K24HL124366 awarded to SP.

---

## Conflicts of Interest

SP is a paid scientific advisor for Fitbit and has been a paid consultant for WW.

---

### Multimedia Appendix 1

Proportion of participants using their assigned app each week of the 12-week intervention, by treatment condition.

[[DOCX File , 43 KB - mhealth\\_v9i4e24249\\_app1.docx](#) ]

---

### Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1382 KB - mhealth\\_v9i4e24249\\_app2.pdf](#) ]

## References

1. Overweight and obesity. National Heart Lung and Blood Institute. URL: <https://www.nhlbi.nih.gov/health-topics/overweight-and-obesity> [accessed 2018-02-15]
2. Kahan S, Manson JE. Nutrition counseling in clinical practice: how clinicians can do better. *J Am Med Assoc* 2017 Sep 26;318(12):1101-1102. [doi: [10.1001/jama.2017.10434](https://doi.org/10.1001/jama.2017.10434)] [Medline: [28880975](https://pubmed.ncbi.nlm.nih.gov/28880975/)]
3. An R, Ji M, Zhang S. Effectiveness of social media-based interventions on weight-related behaviors and body weight status: review and meta-analysis. *Am J Health Behav* 2017 Nov 01;41(6):670-682. [doi: [10.5993/AJHB.41.6.1](https://doi.org/10.5993/AJHB.41.6.1)] [Medline: [29025495](https://pubmed.ncbi.nlm.nih.gov/29025495/)]
4. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. *J Am Diet Assoc* 2011 Jan;111(1):92-102 [FREE Full text] [doi: [10.1016/j.jada.2010.10.008](https://doi.org/10.1016/j.jada.2010.10.008)] [Medline: [21185970](https://pubmed.ncbi.nlm.nih.gov/21185970/)]
5. Yu Z, Sealey-Potts C, Rodriguez J. Dietary self-monitoring in weight management: current evidence on efficacy and adherence. *J Acad Nutr Diet* 2015 Dec;115(12):1931-1938. [doi: [10.1016/j.jand.2015.04.005](https://doi.org/10.1016/j.jand.2015.04.005)] [Medline: [26028176](https://pubmed.ncbi.nlm.nih.gov/26028176/)]
6. Patel ML, Brooks TL, Bennett GG. Consistent self-monitoring in a commercial app-based intervention for weight loss: results from a randomized trial. *J Behav Med* 2020 Jun;43(3):391-401. [doi: [10.1007/s10865-019-00091-8](https://doi.org/10.1007/s10865-019-00091-8)] [Medline: [31396820](https://pubmed.ncbi.nlm.nih.gov/31396820/)]
7. Rogers JM, Ferrari M, Mosely K, Lang CP, Brennan L. Mindfulness-based interventions for adults who are overweight or obese: a meta-analysis of physical and psychological health outcomes. *Obes Rev* 2017 Jan;18(1):51-67. [doi: [10.1111/obr.12461](https://doi.org/10.1111/obr.12461)] [Medline: [27862826](https://pubmed.ncbi.nlm.nih.gov/27862826/)]
8. Pack wraps. University of Massachusetts Medical School. 2006. URL: <https://www.umassmed.edu/globalassets/center-for-tobacco-cttrt/member-area-files/module-five/pack-wraps.pdf> [accessed 2020-02-02]
9. Corwin RL, Hajnal A. Too much of a good thing: neurobiology of non-homeostatic eating and drug abuse. *Physiol Behav* 2005 Sep 15;86(1-2):5-8 [FREE Full text] [doi: [10.1016/j.physbeh.2005.06.021](https://doi.org/10.1016/j.physbeh.2005.06.021)] [Medline: [16081112](https://pubmed.ncbi.nlm.nih.gov/16081112/)]
10. Forman EM, Goldstein SP, Zhang F, Evans BC, Manasse SM, Butryn ML, et al. OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Transl Behav Med* 2019 Mar 01;9(2):236-245 [FREE Full text] [doi: [10.1093/tbm/iby016](https://doi.org/10.1093/tbm/iby016)] [Medline: [29617911](https://pubmed.ncbi.nlm.nih.gov/29617911/)]
11. Carels RA, Hoffman J, Collins A, Raber AC, Cacciapaglia H, O'Brien WH. Ecological momentary assessment of temptation and lapse in dieting. *Eat Behav* 2001;2(4):307-321. [doi: [10.1016/s1471-0153\(01\)00037-x](https://doi.org/10.1016/s1471-0153(01)00037-x)] [Medline: [15001025](https://pubmed.ncbi.nlm.nih.gov/15001025/)]
12. Carels RA, Douglass OM, Cacciapaglia HM, O'Brien WH. An ecological momentary assessment of relapse crises in dieting. *J Consult Clin Psychol* 2004 Apr;72(2):341-348. [doi: [10.1037/0022-006X.72.2.341](https://doi.org/10.1037/0022-006X.72.2.341)] [Medline: [15065966](https://pubmed.ncbi.nlm.nih.gov/15065966/)]
13. McKee HC, Ntoumanis N, Taylor IM. An ecological momentary assessment of lapse occurrences in dieters. *Ann Behav Med* 2014 Dec;48(3):300-310. [doi: [10.1007/s12160-014-9594-y](https://doi.org/10.1007/s12160-014-9594-y)] [Medline: [24562984](https://pubmed.ncbi.nlm.nih.gov/24562984/)]
14. Grilo CM, Shiffman S, Wing RR. Relapse crises and coping among dieters. *J Consult Clin Psychol* 1989 Aug;57(4):488-495. [doi: [10.1037//0022-006x.57.4.488](https://doi.org/10.1037//0022-006x.57.4.488)] [Medline: [2768608](https://pubmed.ncbi.nlm.nih.gov/2768608/)]
15. Forman EM, Goldstein SP, Crochiere RJ, Butryn ML, Juarascio AS, Zhang F, et al. Randomized controlled trial of OnTrack, a just-in-time adaptive intervention designed to enhance weight loss. *Transl Behav Med* 2019 Nov 25;9(6):989-1001. [doi: [10.1093/tbm/ibz137](https://doi.org/10.1093/tbm/ibz137)] [Medline: [31602471](https://pubmed.ncbi.nlm.nih.gov/31602471/)]
16. Tulu B, Ruiz C, Allard J, Acheson J, Busch A, Roskuski A. Slip buddy: a mobile health intervention to prevent overeating. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; 2017; Manoa, Hawaii. [doi: [10.24251/hicss.2017.436](https://doi.org/10.24251/hicss.2017.436)]
17. Pagoto S, Waring M, Olendzki E, Oleski J, May C, Evans M. The feasibility of incentivizing participation in an online social network weight loss program. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; 2017; Hawaii. [doi: [10.24251/hicss.2017.457](https://doi.org/10.24251/hicss.2017.457)]
18. Waring ME, Simas TAM, Oleski J, Xiao RS, Mulcahy JA, May CN, et al. Feasibility and acceptability of delivering a postpartum weight loss intervention via facebook: a pilot study. *J Nutr Educ Behav* 2018 Jan;50(1):70-74 [FREE Full text] [doi: [10.1016/j.jneb.2017.09.025](https://doi.org/10.1016/j.jneb.2017.09.025)] [Medline: [29325665](https://pubmed.ncbi.nlm.nih.gov/29325665/)]
19. Pagoto S, Waring M, Olendzki E, Oleski J, May C, Evans M. The feasibility of incentivizing participation in an online social network weight loss program. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; 2017; Hawaii.
20. Payne JE, Turk MT, Kalarchian MA, Pellegrini CA. Defining adherence to dietary self-monitoring using a mobile app: a narrative review. *J Acad Nutr Diet* 2018 Nov;118(11):2094-2119. [doi: [10.1016/j.jand.2018.05.011](https://doi.org/10.1016/j.jand.2018.05.011)] [Medline: [30115555](https://pubmed.ncbi.nlm.nih.gov/30115555/)]
21. Turner-McGrievy GM, Dunn CG, Wilcox S, Boutté AK, Hutto B, Hoover A, et al. Defining adherence to mobile dietary self-monitoring and assessing tracking over time: tracking at least two eating occasions per day is best marker of adherence within two different mobile health randomized weight loss interventions. *J Acad Nutr Diet* 2019 Sep;119(9):1516-1524 [FREE Full text] [doi: [10.1016/j.jand.2019.03.012](https://doi.org/10.1016/j.jand.2019.03.012)] [Medline: [31155473](https://pubmed.ncbi.nlm.nih.gov/31155473/)]
22. Brooke J. SUS-A quick and dirty usability scale. In: Usability evaluation in industry. Florida: CRC Press; 1996:194.
23. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Studies*. 2009. URL: <https://dl.acm.org/doi/10.5555/2835587.2835589> [accessed 2021-02-27]
24. Brooke J. SUS: a retrospective. *J Usability Stud*. 2013. URL: <https://uxpajournal.org/sus-a-retrospective/> [accessed 2021-02-27]

25. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res* 2011 May;45(5):626-629 [FREE Full text] [doi: [10.1016/j.jpsychires.2010.10.008](https://doi.org/10.1016/j.jpsychires.2010.10.008)] [Medline: [21035130](https://pubmed.ncbi.nlm.nih.gov/21035130/)]
26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
27. 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](https://doi.org/10.20982/tqmp.08.1.p023)] [Medline: [22833776](https://pubmed.ncbi.nlm.nih.gov/22833776/)]
28. Dunn CG, Turner-McGrievy GM, Wilcox S, Hutto B. Dietary self-monitoring through calorie tracking but not through a digital photography app is associated with significant weight loss: the 2SMART pilot study - a 6-month randomized trial. *J Acad Nutr Diet* 2019 Sep;19(9):1525-1532. [doi: [10.1016/j.jand.2019.03.013](https://doi.org/10.1016/j.jand.2019.03.013)] [Medline: [31155474](https://pubmed.ncbi.nlm.nih.gov/31155474/)]
29. Daim TU, Basoglu N, Topacan U. Adoption of health information technologies: the case of a wireless monitor for diabetes and obesity patients. *Technol Anal Strateg Manag* 2013 Sep;25(8):923-938. [doi: [10.1080/09537325.2013.823150](https://doi.org/10.1080/09537325.2013.823150)]
30. Yu P, Li H, Gagnon MP. Health IT acceptance factors in long-term care facilities: a cross-sectional survey. *Int J Med Inform* 2009 Apr;78(4):219-229 [FREE Full text] [doi: [10.1016/j.ijmedinf.2008.07.006](https://doi.org/10.1016/j.ijmedinf.2008.07.006)] [Medline: [18768345](https://pubmed.ncbi.nlm.nih.gov/18768345/)]
31. Forman EM, Schumacher LM, Crosby R, Manasse SM, Goldstein SP, Butryn ML, et al. Ecological momentary assessment of dietary lapses across behavioral weight loss treatment: characteristics, predictors, and relationships with weight change. *Ann Behav Med* 2017 Oct;51(5):741-753 [FREE Full text] [doi: [10.1007/s12160-017-9897-x](https://doi.org/10.1007/s12160-017-9897-x)] [Medline: [28281136](https://pubmed.ncbi.nlm.nih.gov/28281136/)]
32. Evers C, Dingemans A, Junghans AF, Boevé A. Feeling bad or feeling good, does emotion affect your consumption of food? A meta-analysis of the experimental evidence. *Neurosci Biobehav Rev* 2018 Sep;92:195-208. [doi: [10.1016/j.neubiorev.2018.05.028](https://doi.org/10.1016/j.neubiorev.2018.05.028)] [Medline: [29860103](https://pubmed.ncbi.nlm.nih.gov/29860103/)]
33. Goldstein SP, Thomas JG, Foster GD, Turner-McGrievy G, Butryn ML, Herbert JD, et al. Refining an algorithm-powered just-in-time adaptive weight control intervention: a randomized controlled trial evaluating model performance and behavioral outcomes. *Health Informatics J* 2020 Dec 06;26(4):2315-2331 [FREE Full text] [doi: [10.1177/1460458220902330](https://doi.org/10.1177/1460458220902330)] [Medline: [32026745](https://pubmed.ncbi.nlm.nih.gov/32026745/)]
34. Turner-McGrievy GM, Campbell MK, Tate DF, Truesdale KP, Bowling JM, Crosby L. Pounds off digitally study: a randomized podcasting weight-loss intervention. *Am J Prev Med* 2009 Oct;37(4):263-269 [FREE Full text] [doi: [10.1016/j.amepre.2009.06.010](https://doi.org/10.1016/j.amepre.2009.06.010)] [Medline: [19765496](https://pubmed.ncbi.nlm.nih.gov/19765496/)]
35. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. *J Med Internet Res* 2015 Oct 23;17(10):e240 [FREE Full text] [doi: [10.2196/jmir.4897](https://doi.org/10.2196/jmir.4897)] [Medline: [26499966](https://pubmed.ncbi.nlm.nih.gov/26499966/)]
36. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. *Obesity (Silver Spring)* 2012 Jun;20(6):1234-1239 [FREE Full text] [doi: [10.1038/oby.2011.140](https://doi.org/10.1038/oby.2011.140)] [Medline: [21633403](https://pubmed.ncbi.nlm.nih.gov/21633403/)]
37. Joiner KL, Nam S, Whittemore R. Lifestyle interventions based on the diabetes prevention program delivered via eHealth: a systematic review and meta-analysis. *Prev Med* 2017 Jul;100:194-207 [FREE Full text] [doi: [10.1016/j.ypmed.2017.04.033](https://doi.org/10.1016/j.ypmed.2017.04.033)] [Medline: [28456513](https://pubmed.ncbi.nlm.nih.gov/28456513/)]

## Abbreviations

- DPP:** Diabetes Prevention Program  
**IRR:** interrater reliability  
**SUS:** System Usability Scale  
**WW:** Weight Watchers

*Edited by L Buis; submitted 15.09.20; peer-reviewed by S Goldstein, L Schumacher, K Joiner; comments to author 03.11.20; revised version received 29.12.20; accepted 22.02.21; published 01.04.21.*

### *Please cite as:*

Pagoto S, Tulu B, Waring ME, Goetz J, Bibeau J, Divito J, Groshon L, Schroeder M  
Slip Buddy App for Weight Management: Randomized Feasibility Trial of a Dietary Lapse Tracking App  
*JMIR Mhealth Uhealth* 2021;9(4):e24249  
URL: <https://mhealth.jmir.org/2021/4/e24249>  
doi: [10.2196/24249](https://doi.org/10.2196/24249)  
PMID: [33792547](https://pubmed.ncbi.nlm.nih.gov/33792547/)

©Sherry Pagoto, Bengisu Tulu, Molly E Waring, Jared Goetz, Jessica Bibeau, Joseph Divito, Laurie Groshon, Matthew Schroeder. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 01.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Quality, Features, and Presence of Behavior Change Techniques in Mobile Apps Designed to Improve Physical Activity in Pregnant Women: Systematic Search and Content Analysis

Melanie Hayman<sup>1</sup>, PhD; Kristie-Lee Alfrey<sup>1</sup>, BPsychSc(Hons); Summer Cannon<sup>1</sup>, BPsychSc(Hons); Stephanie Alley<sup>1</sup>, PhD; Amanda L Rebar<sup>1</sup>, PhD; Susan Williams<sup>1</sup>, PhD; Camille E Short<sup>2</sup>, PhD; Abby Altazan<sup>3</sup>, MSc; Natalie Comardelle<sup>3</sup>, BSc; Sinead Currie<sup>4</sup>, PhD; Caitlin Denton<sup>3</sup>, BSc; Cheryce L Harrison<sup>5</sup>, PhD; Tayla Lamerton<sup>6</sup>, PhD; Gabriela P Mena<sup>6</sup>, MD; Lisa Moran<sup>5</sup>, PhD; Michelle Mottola<sup>7</sup>, PhD; Taniya S Nagpal<sup>8</sup>; Lisa Vincze<sup>9</sup>, PhD; Stephanie Schoeppe<sup>1</sup>, PhD

<sup>1</sup>School of Health, Medical and Applied Sciences, CQUniversity, Rockhampton, Australia

<sup>2</sup>School of Psychological Sciences, University of Melbourne, Melbourne, Australia

<sup>3</sup>Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, LA, United States

<sup>4</sup>Division of Psychology, Stirling University, Scotland, United Kingdom

<sup>5</sup>School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

<sup>6</sup>School of Human Movement and Nutrition Sciences, University of Queensland, Brisbane, Australia

<sup>7</sup>Faculty of Health Sciences in Kinesiology, University of Western Ontario, London, ON, Canada

<sup>8</sup>School of Human Kinetics, University of Ottawa, Ottawa, ON, Canada

<sup>9</sup>School of Allied Health Sciences - Nutrition and Dietetics, Griffith University, Gold Coast, Australia

**Corresponding Author:**

Melanie Hayman, PhD

School of Health, Medical and Applied Sciences

CQUniversity

Bruce Highway

Rockhampton, 4701

Australia

Phone: 61 49306912 ext 56912

Email: [m.j.hayman@cqu.edu.au](mailto:m.j.hayman@cqu.edu.au)

## Abstract

**Background:** Physical activity during pregnancy is associated with several health benefits for the mother and child. However, very few women participate in regular physical activity during pregnancy. eHealth platforms (internet and mobile apps) have become an important information source for pregnant women. Although the use of pregnancy-related apps has significantly increased among pregnant women, very little is known about their theoretical underpinnings, including their utilization of behavior change techniques (BCTs). This is despite research suggesting that inclusion of BCTs in eHealth interventions are important for promoting healthy behaviors, including physical activity.

**Objective:** The aim of this study was to conduct a systematic search and content analysis of app quality, features, and the presence of BCTs in apps designed to promote physical activity among pregnant women.

**Methods:** A systematic search in the Australian App Store and Google Play store using search terms relating to *exercise* and *pregnancy* was performed. App quality and features were assessed using the 19-item Mobile App Rating Scale (MARS), and a taxonomy of BCTs was used to determine the presence of BCTs (26 items). BCTs previously demonstrating efficacy in behavior changes during pregnancy were also identified from a literature review. Spearman correlations were used to investigate the relationships between app quality, app features, and number of BCTs identified.

**Results:** Nineteen exercise apps were deemed eligible for this review and they were accessed via Google Play (n=13) or App Store (n=6). The MARS overall quality scores indicated moderate app quality (mean 3.5 [SD 0.52]). Functionality was the highest scoring MARS domain (mean 4.2 [SD 0.5]), followed by aesthetics (mean 3.7 [SD 0.6]) and information quality (mean 3.16 [SD 0.42]). Subjective app quality (mean 2.54 [SD 0.64]) and likelihood for behavioral impact (mean 2.5 [SD 0.6]) were the lowest scoring MARS domains. All 19 apps were found to incorporate at least two BCTs (mean 4.74, SD 2.51; range 2-10). However, only 11 apps included BCTs that previously demonstrated efficacy for behavior change during pregnancy, the most common

being *provide opportunities for social comparison* (n=8) and *prompt self-monitoring of behavior* (n=7). There was a significant positive correlation between the number of BCTs with engagement and aesthetics scores, but the number of BCTs was not significantly correlated with functionality, information quality, total MARS quality, or subjective quality.

**Conclusions:** Our findings showed that apps designed to promote physical activity among pregnant women were functional and aesthetically pleasing, with overall moderate quality. However, the incorporation of BCTs was low, with limited prevalence of BCTs previously demonstrating efficacy in behavior change during pregnancy. Future app development should identify and adopt factors that enhance and encourage user engagement, including the use of BCTs, especially those that have demonstrated efficacy for promoting physical activity behavior change among pregnant women.

(*JMIR Mhealth Uhealth* 2021;9(4):e23649) doi:[10.2196/23649](https://doi.org/10.2196/23649)

## KEYWORDS

pregnancy; exercise; physical activity; mobile health (mHealth); applications; MARS; behavior change techniques; mobile phone

## Introduction

Physical activity during pregnancy is associated with a variety of health benefits, including reduced risk of excessive gestational weight gain, gestational diabetes, gestational hypertension, preeclampsia, the severity of pelvic girdle pain, macrosomia, instrumental delivery, postpartum weight retention, urinary incontinence, and depressive disorders [1,2]. Despite the many health benefits of physical activity during pregnancy, few women participate in regular physical activity during pregnancy [3]. In addition, women tend to reduce or cease their participation in physical activity once they become pregnant and throughout the course of their pregnancy [3,4]. This may be as a result of various barriers, including mother-child safety concerns, fatigue, change in body shape, and associated pain. Further, a lack of provision of adequate information, knowledge, social support, and self-efficacy for behavior change are other issues that may exacerbate the decline in activity levels throughout pregnancy [2,5].

There are many avenues for women to access information and support related to maintaining a healthy pregnancy, including information about physical activity behaviors. Historically, pregnant women have accessed information from doctors, midwives, family, and friends to guide and inform their physical activity behaviors. However, eHealth platforms such as the internet and mobile apps are now altering the way women access this information [6] and they have become an important information source for pregnant women [7,8]. In fact, a recent Australian study among 410 pregnant women investigated the use of pregnancy and parenting apps and found that almost three-quarters of the studied women used at least one of these types of pregnancy apps [9]. In addition, more than half of the participants reported using 2-4 apps throughout their pregnancy. The frequency of app use was also significant, with almost a quarter of pregnant women reporting daily use of apps [9]. While the use of pregnancy-related apps has significantly increased among pregnant women [8], very little is known about their theoretical underpinnings, including their utilization of behavior change techniques (BCTs). This is despite research suggesting that inclusion of BCTs in eHealth interventions can play an important role in improving, supporting, and maintaining healthy behaviors, including physical activity [10,11].

In 2013, Currie et al [12] systematically evaluated the content of physical activity interventions designed to reduce the decline

of physical activity in pregnant women with a specific emphasis on BCTs [13] employed to elicit this change. Six common BCTs shown to have some efficacy in improving physical activity behaviors were identified: prompt intention formation, prompt specific goal setting, prompt review of behavioral goals, prompt self-monitoring of behavior, provide feedback on performance, and provide opportunities for social comparison [12]. Since this review, many behavior change interventions have used these BCTs to promote positive physical activity behaviors among pregnant women [14].

Previous reviews of physical activity apps for other population groups suggest that commercial apps often lack evidence-based BCTs that have demonstrated efficacy for encouraging physical activity behavior change [15-17]. However, no such review of commercial apps designed to promote physical activity among pregnant women has been conducted. Thus, the appropriateness of these apps to promote physical activity during pregnancy is unknown. This review aimed to systematically evaluate the appropriateness of the apps designed to promote physical activity among pregnant women by using a systematic search and content analysis. Apps available through the Australian App Store and Google Play stores were accessed using the MARS tool for app quality and features. A taxonomy of BCT was also used to assess the presence of BCTs utilized within the apps, including BCTs that have demonstrated efficacy for promoting physical activity behavior change among pregnant women.

## Methods

### Methodological Approach

The methodological approach used in this study was informed by previous app reviews. These reviews explore app quality, features, and BCTs among apps designed to (1) improve diet, physical activity, and sedentary behavior in children and adolescents [15] and (2) provide nutritional advice to pregnant women [18].

### Search Strategy

Systematic searches were conducted in the Australian App Store and Google Play stores between October 2018 and February 2019. Apps were identified using systematic combinations of the following search terms: pregnancy, pregnant, prenatal, postnatal, exercises, exercise, fitness, workout, and physical

activity. These search term combinations were entered individually in the App Store and Google Play databases without any specified search categories, and search results were ordered by relevance (see [Multimedia Appendix 1](#)). As Google Play search results were capped at 250 apps, only the title and description of the first 250 relevant apps (in Google Play and App Store) were screened.

### Inclusion Criteria and Selection Process

Apps were considered for inclusion if the description of the app in the stores specified pregnancy content and physical activity or exercise. Apps were included if they (1) targeted pregnant women, (2) had a focus on physical activity or exercise, (3) were available in English, and (4) had a user rating of at least 4.5 (scale range 1-5) in either of the stores (as similarly done elsewhere [17]) as a measure of app popularity. Both free and paid apps were eligible for inclusion; however, apps requiring external devices (eg, Kegel device, activity monitor, hardcopy books) were excluded due to limitations regarding device access and use. App selection and assessments were undertaken between October 2018 and April 2019.

As per best practice for systematic reviews [19], 2 reviewers (KLA and SC) independently reviewed the titles, images, and descriptions of each identified app for inclusion in the review. Disagreement was resolved by discussion and consensus with a third reviewer (MH). Each of the eligible apps were examined independently by 2 of the 18 reviewers (ie, the authors), who were recognized as having expertise in behavior change or physical activity during pregnancy. If there was any notable disagreement in variance (eg, disagree versus agree) among the app assessment scores, a third reviewer would also be assigned. Each reviewer was allocated 4-6 apps to examine, determined by device accessibility (Apple or Android). Examination of apps included downloading, user testing, and assessing app features and quality criteria. Each app was allocated to an expert in behavior change and to an expert in physical activity during pregnancy for review. Incorporation of BCTs within each app were independently reviewed by 2 reviewers (KLA and SC). Any disagreements/discrepancies between reviewers KLA and SC were resolved by consultation with a third reviewer (MC). If an app was available in both App Store and Google Play, either version could be utilized for testing, regardless of differences in app user ratings. To maintain a consistent cost status and baseline assessment, if an app offered a free version and a paid version, the free version was included. To maintain this consistency, freemium content (ie, extra content at a cost) was not accessed and apps requiring paid subscriptions were excluded.

### Data Extraction

Data extraction was conducted using a standard information spreadsheet and the Mobile App Rating Scale (MARS) [20]. Similar methods have been utilized in previous app reviews [15,21]. For all included apps, app name, developer, version, store (App Store, Google Play), cost (free, paid), average user rating (at least 4.5+), MARS focus points (what the app targets, eg, increase happiness/well-being, behavior change, entertainment, physical health), MARS theoretical background/strategies (eg, assessment, information/education,

goal setting, advice/tips/strategies/skills training), and MARS technical aspects (eg, allows sharing, allows password protection, sends reminders) were extracted (see [Multimedia Appendix 2](#) and [Multimedia Appendix 3](#)).

### App Features and Quality Assessment

App features and quality were assessed using the MARS [20], as per prior app reviews [15,22]. The MARS consists of 19 items grouped in 4 domains: engagement (entertainment, interest, customization, interactivity, and target group); functionality (performance, ease of use, navigation, and gestural design); aesthetics (layout, graphics, and visual appeal); and information quality (accuracy of app description, goals, quality, and quantity of information, visual information, credibility, and evidence base). Additional MARS domains of subjective app quality (recommendation, potential use, payment, and overall rating) and likelihood of behavioral impact (awareness, knowledge, attitudes, intention to change, help seeking, and behavior change) were also included. Items were measured on a 5-point scale (1=inadequate to 5=excellent) and a score for each domain was computed as the mean of the items in that domain; the overall score was computed as an average across the domains [20]. Final scores for each app and MARS items were calculated using the means of the reviewer scores (see [Multimedia Appendix 4](#)).

### BCT Identification

The assessment of the presence or absence of BCTs for improving physical activity behavior was guided by the taxonomy of BCTs developed by Abraham and Michie [13]. A dichotomous score of 0 (absent) or 1 (present) was applied for each of the 26 BCTs, resulting in a total score of 0-26 (see [Multimedia Appendix 5](#)). This approach has been applied in similar app reviews and content analyses [15,23,24].

### Identification of Evidence-Based BCTs

A brief literature search was employed to understand the BCTs that may be effective in supporting behavior change during pregnancy. A systematic review by Currie et al [12] described 6 BCTs that hold efficacy in reducing the decline of physical activity among pregnant women. These BCTs include prompt intention formation, prompt specific goal setting, prompt review of behavioral goals, prompt self-monitoring of behavior, provide feedback on performance, and provide opportunities for social comparison. As pregnancy-effective BCTs, these were specifically highlighted during analysis and results.

### Statistical Analyses

In addition to descriptive statistics (mean, standard deviation, and range) calculated for each of the 6 MARS domains, frequencies (numbers and percentages) of each of the 26 BCTs included in the apps were calculated. Krippendorff's alpha ( $K\alpha$ ) was used to evaluate interrater reliability for the app quality assessment and the presence of BCTs within the apps [25]. Spearman correlations were used to examine the relationships between app quality, number of technical app features, and number of BCTs incorporated in the apps. All statistical analyses were conducted using SPSS Statistics version 26.0 (IBM Corp) with significance levels set at  $P < .05$ .



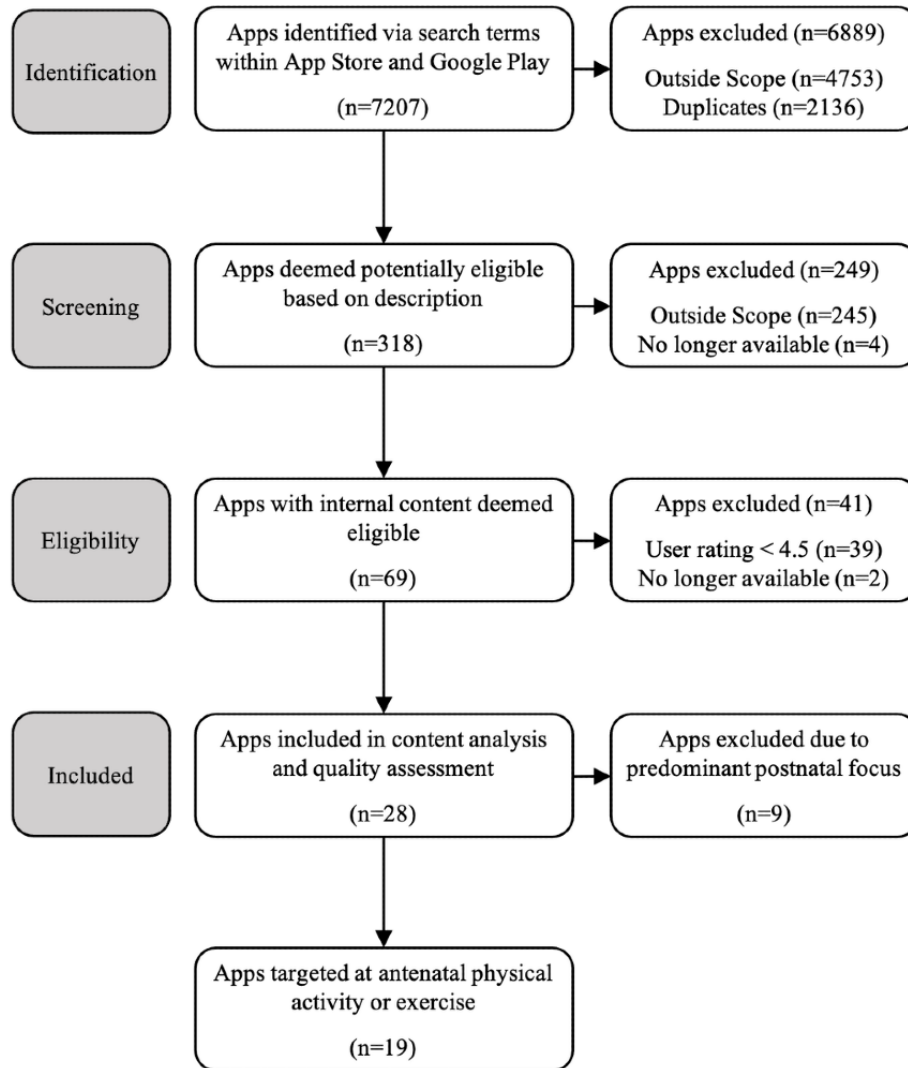
## Results

### App Selection

A flowchart of the app selection process is presented in [Figure 1](#). A total of 7207 apps were identified and screened in the App Store and Google Play. Of these, 318 apps were further screened

by description and 69 apps held content considered eligible for inclusion. The user rating criteria of 4.5+ was applied and apps found to focus solely on postnatal physical activity/exercise were omitted. A total of 19 apps targeting physical activity during pregnancy were included in the content analysis and quality assessment.

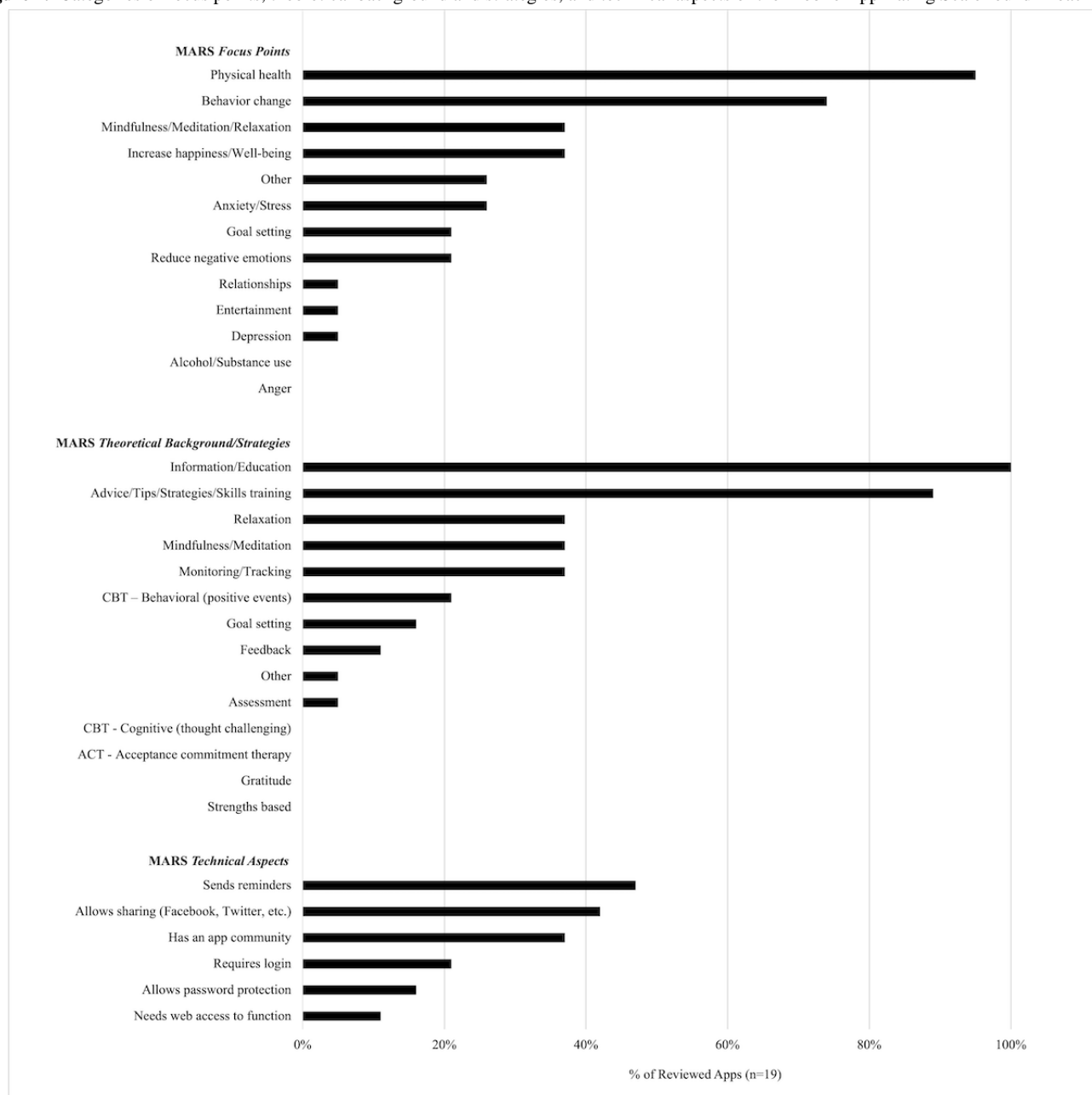
**Figure 1.** PRISMA flow chart of the app selection process.



### App Characteristics

Of the 19 reviewed antenatal physical activity apps, 13 were accessed via Google Play and 6 were accessed via App Store (see [Multimedia Appendix 2](#)). Apps were free to download, with the exception of one. The average star rating for the apps was 4.69 (SD 0.22), with a wide range of the number of users

rating each app (mean 1875.16, SD 3549.82; range 1-13,000). On average, the 19 apps were found to contain few MARS-related categories: MARS focus points (mean 3.53 [SD 1.90]), MARS theoretical background/strategies (mean 3.58 [SD 1.98]), and MARS technical aspects (mean 1.74 [SD 1.73]). [Figure 2](#) and [Multimedia Appendix 3](#) detail the MARS categories for each of the 19 apps.

**Figure 2.** Categories of focus points, theoretical background and strategies, and technical aspects of the Mobile App Rating Scale found in each app.

## App Quality

The average MARS overall quality score was 3.5 out of 5 with a range of 2.4-4.3, which was considered to be of moderate quality. Functionality was the highest scoring domain (mean 4.2 [SD 0.5]), followed by aesthetics (mean 3.7 [0.6]), information quality (mean 3.16 [SD 0.42]), and engagement (mean 3.01 [SD 0.9]). Subjective app quality (mean 2.5 [SD 0.6]) and likelihood for behavioral impact (mean 2.5 [SD 0.6])

were equally the lowest scoring MARS domains. [Table 1](#) provides a summary of the MARS scores. A detailed summary of the quality assessment of the included apps is presented in [Multimedia Appendix 4](#). Interrater reliability [25] for app quality resulted in low reliability (mean  $K\alpha$  0.3, SD 0.37). However, there was no notable disagreement in variance (eg, disagree versus agree) among the app assessment scores; thus, a third reviewer was not required for any further app assessment.

**Table 1.** Summary of the Mobile App Rating Scale (scale 1-5) scores across the 19 reviewed apps.<sup>a</sup>

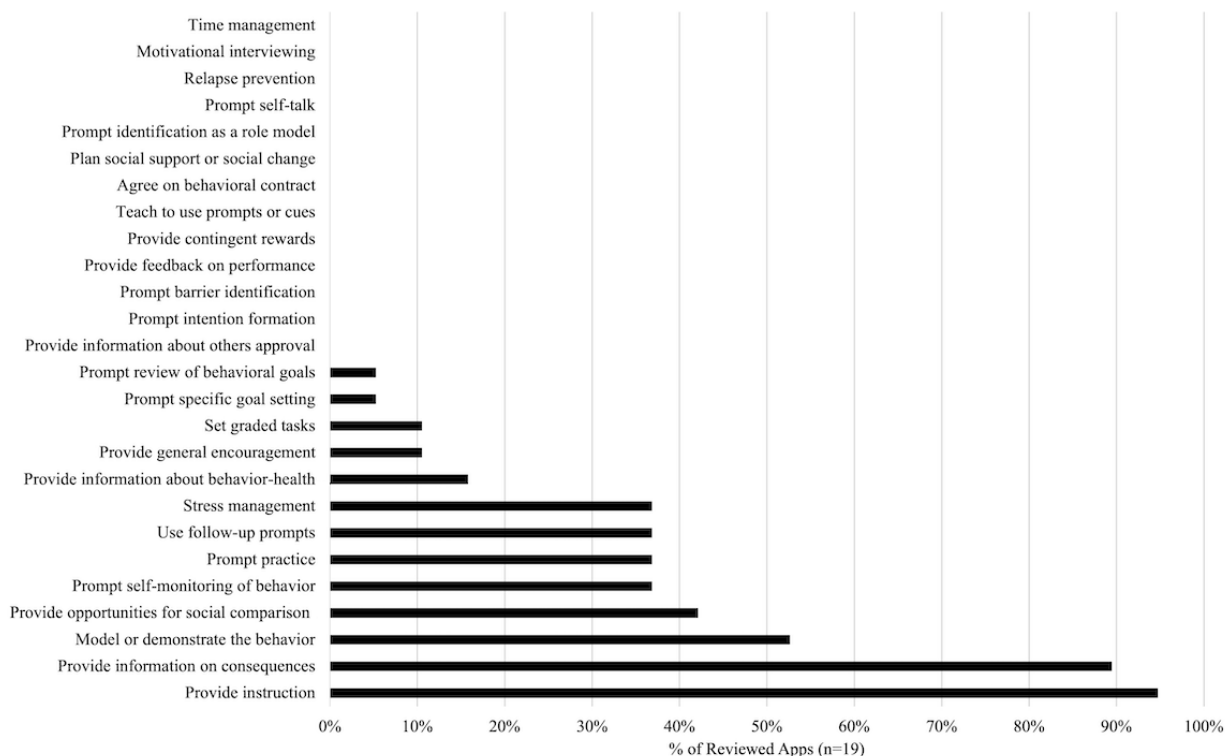
Mobile App Rating Scale domain	Mean score (SD)	Median score (IQR)	Range of scores (min-max)
Functionality	4.22 (0.49)	4.3 (0.75)	1.7 (3.3-5)
Aesthetics	3.69 (0.64)	4 (0.95)	2.2 (2.3-4.5)
Overall quality	3.52 (0.52)	3.5 (0.7)	1.9 (2.4-4.3)
Information quality	3.19 (0.42)	3.2 (0.75)	1.4 (2.6-4)
Engagement	3.01 (0.9)	3 (1.4)	3.3 (1.2-4.5)
Subjective quality	2.54 (0.64)	2.5 (0.95)	2.3 (1.5-3.8)
Likelihood of behavioral impact	2.54 (0.62)	2.6 (0.6)	2.2 (1.6-3.8)

<sup>a</sup>Score 1=inadequate; score 5=excellent.

### Presence of BCTs

The presence and types of BCTs found within the reviewed apps are presented in [Figure 3](#) and [Multimedia Appendix 5](#). Interrater reliability for evaluating the presence of BCTs in the apps was high ( $K\alpha$  0.85, percent agreement 95%). All reviewed apps incorporated at least two BCTs. Commonly included BCTs included *provide instructions* (18/19, 95%), *provide information*

*on consequences* (17/19, 89%), *model or demonstrate the behavior* (10/19, 53%), and *provide opportunities for social comparison* (8/19, 42%). The average number of BCTs per app was 4.74 (SD 2.51) (range 2-10). Apps with the highest number of BCTs included were Pregnancy Week by Week Tracker (10 BCTs), iMum-Pregnancy & Fertility (9 BCTs), and Pregnancy Tracker & Countdown (9 BCTs).

**Figure 3.** Presence and types of behaviour change techniques found in the selected apps.

### Presence of Evidence-Based BCTs

Currie et al [12] identified 6 common BCTs shown to have some efficacy in reducing the decline in physical activity behaviors among pregnant women, namely, *prompt intention formation*, *prompt specific goal setting*, *prompt review of behavioral goals*, *prompt self-monitoring of behavior*, *provide*

*feedback on performance*, and *provide opportunities for social comparison*. Of the 19 apps reviewed in the present study, 11 apps contained at least one of these evidence-based BCTs (range 1-3) and 4 contained more than one of these evidence-based BCTs. [Table 2](#) details the evidence-based BCTs included in each of the 11 apps.

**Table 2.** Apps containing evidence-based techniques for behavior change during pregnancy.<sup>a</sup>

App name	Evidence-based behavior change techniques					
	Prompt intention formation (n=0)	Prompt specific goal setting (n=1)	Prompt review of behavioral goals (n=1)	Prompt self-monitoring of behavior (n=7)	Provide feedback on performance (n=0)	Provide opportunities for social comparison (n=8)
Pregnancy+ (n=3)			✓	✓		✓
Pregnancy Week by Week Tracker (n=3)		✓		✓		✓
Pregnancy Tracker & Countdown (n=2)				✓		✓
Pregnancy Work-outs-Baby2Body (n=2)				✓		✓
9Months Guide (n=1)						✓
Get Parenting Pregnancy Tips. Moms Pregnancy App (n=1)						✓
I'm Pregnant-Pregnancy Tracker (n=1)				✓		
iMum-Pregnancy & Fertility (n=1)				✓		
Kegel Exercises (n=1)				✓		
Pregnancy Guide (n=1)						✓
Pregnancy Health (n=1)						✓

<sup>a</sup>Apps not identified as containing a key behavior change technique are not noted in the table.

### Relationships Between App Quality, App Features, and BCTs

Spearman correlations between the MARS overall quality, number of MARS focus points, number of MARS theoretical background/strategies, number of MARS technical aspects used in the app, MARS subjective quality, and the number of BCTs are presented in Table 3. The number of identified BCTs was positively associated with the MARS engagement score ( $\rho=0.55$ ,  $P=.01$ ) and aesthetics score ( $\rho=0.46$ ,  $P=.046$ ). MARS functionality, information quality, total MARS quality, and

subjective quality were not significantly correlated with the number of BCTs. The number of app focus points was positively associated with the MARS engagement score ( $\rho=0.58$ ,  $P=.009$ ), aesthetics score ( $\rho=0.58$ ,  $P=.008$ ), total MARS quality score ( $\rho=0.55$ ,  $P=.02$ ), and subjective quality score ( $\rho=0.46$ ,  $P=.048$ ). The number of technical aspects within apps was positively correlated with the MARS engagement score ( $\rho=0.63$ ,  $P=.004$ ) but not with any of the other MARS scores. Further, the MARS subjective quality scores were positively correlated with all other quality subscores (Table 3) and the total MARS quality score ( $\rho=0.90$ ,  $P<.001$ ).

**Table 3.** Correlations between Mobile App Rating Scale (MARS) overall quality, number of MARS focus points, number of MARS theoretical background/strategies, number of technical aspects, and MARS subjective quality.

MARS subscales	Behavior change techniques	MARS focus points	MARS theoretical background/strategies	MARS technical aspects	MARS subjective quality
MARS engagement	0.55 <sup>a</sup>	0.58 <sup>b</sup>	0.69 <sup>b</sup>	0.63 <sup>b</sup>	0.69 <sup>b</sup>
MARS functionality	0.02	0.30	0.36	-0.24	0.64 <sup>b</sup>
MARS aesthetics	0.46 <sup>a</sup>	0.59 <sup>b</sup>	0.61 <sup>b</sup>	0.37	0.90 <sup>b</sup>
MARS information quality	0.32	0.34	0.37	0.10	0.69 <sup>b</sup>
MARS overall quality	0.45	0.55 <sup>a</sup>	0.67 <sup>b</sup>	0.31	0.90 <sup>b</sup>
MARS subjective quality	0.42	0.46 <sup>a</sup>	0.60 <sup>b</sup>	0.36	— <sup>c</sup>

<sup>a</sup>Correlation was significant at  $P < .05$ .

<sup>b</sup>Correlation was significant at  $P < .01$ .

<sup>c</sup>Not applicable.

## Discussion

The primary aim of this app review was to assess app quality and features and presence of BCTs applied within commercial apps promoting physical activity in women during pregnancy. The secondary aim of this app review was to test the relationships between the 6 MARS domains of app quality and the number of MARS technical aspects, theoretical strategies, focus points, and BCTs. In summary, our findings demonstrate moderate MARS overall quality scores, with MARS domains of functionality and aesthetics scoring the highest and the domains of subjective app quality and likelihood for behavioral impact scoring the lowest. An average of 4.74 BCTs per app were used, with the most common BCTs being *provide information on consequences* and *provide instructions*. Most apps had none or one of the BCTs that demonstrated efficacy in promoting physical activity behaviors during pregnancy, with the most common BCTs being *opportunities for social comparison* and *prompt self-monitoring of behavior*.

Many commercially available physical activity apps for pregnant women were identified—with 69 relevant apps and 19 apps remaining after apps targeting postnatal physical activity ( $n=9$ ) and apps with user ratings lower than 4.5 ( $n=39$ ) were excluded. This finding demonstrates a market for apps targeting physical activity during pregnancy, which is not surprising, considering the findings of Lupton and Pedersen [9] who have previously reported that pregnant women frequently use pregnancy-related apps. The 69 apps identified as targeting exercise in pregnancy is comparable to the 51 apps Brown et al [18] identified as targeting nutrition in pregnancy. Pregnancy seems to be a time in women's lives in which they are actively seeking and using apps to support their behaviors, including physical activity.

The high MARS functionality scores and moderate-to-high MARS aesthetics scores may be a consequence of developers focusing on user experience for visually pleasing and user-friendly apps [26]. Although such features are important for attracting users, they do not translate to behavior change [27,28]. There is also room for improvement in app design in relation to the low MARS scores for quality of information and

engagement. Given that pregnant women believe the information in apps has high credibility [9], it is important that apps ensure that evidence-based information is provided. It is also important that apps improve ratings of engagement, as engagement in app content is associated with behavior change [22].

Ratings of engagement could be improved by increasing the use of BCTs. Our study found an association between the number of BCTs and improved MARS engagement and aesthetics scores, which is consistent with the findings from a review investigating apps for weight management in adults [22]. The association between the number of BCTs and the engagement score could be due to the provision of additional content to engage in, increased usefulness, and perceived efficacy. It may be that BCTs are considered as enticements to potential users, thereby increasing engagement. If this is the case, this provides additional incentives for app developers to include BCTs. In line with that reported by Schoeppe et al [15] and Bardus et al [22], this study also found technical aspects of health behavior change apps to be associated with the MARS engagement score. Further, the number of theoretical strategies and the number of focus points were associated with MARS engagement, aesthetics, total quality, and subjective quality scores. Improving the number of technical aspects, theoretical strategies, and focus points may therefore help to further improve engagement of apps designed to promote physical activity among pregnant women.

An average of 4.74 BCTs were used in the reviewed apps. This is in line with the findings from previous research that found apps for physical activity promotion in adults to have an average of 4.2 [16] and 5.0 BCTs [24]. The number of BCTs identified in this review is, however, higher than those identified in a recent review exploring the number of BCTs among apps targeting nutrition behaviors in pregnancy [18], which reported an average of 3 BCTs per app [18]. We know that health behavior interventions using alternative modes of delivery (eg, websites) have improved outcomes when more BCTs are used [29]. Webb et al [29] suggest that this improvement may be a result of using a combination of BCTs that together target several stages and aspects of behavior change. However, we do

not know the optimal number or combination of BCTs necessary to increase physical activity.

The most common BCTs identified in this study were *provide information on consequences* and *provide instructions*. Schoeppe et al [15] also found that providing instructions was a commonly used BCT in physical activity and diet apps for children and adolescents. This finding differs from the most common BCTs identified in apps targeting weight and physical activity in adults, which are goal setting, self-monitoring, and performance feedbacks [15,29]. Unfortunately, *provide information on consequences* and *provide instructions* have not been shown to be effective at improving physical activity behaviors in pregnant women [12] or even in the general adult population [27].

Most apps in this review had none or 1 BCT, which previously demonstrated efficacy in behavior change during pregnancy. Including more of these evidence-based BCTs in the context of pregnancy may improve the ability of the apps to support pregnant women in increasing their physical activity [12]. The evidence-based BCTs of intention formation, goal setting, review of goals, and feedback on performance have been successfully implemented within physical activity apps targeting other population groups [15,16]. This demonstrates that it is feasible to deliver these evidence-based BCTs through an app. In particular, feedback on performance was identified as one of the most commonly used BCTs in apps targeting physical activity, diet, and sedentary behavior in children and adults [15]. Feedback on performance requires the measurement of behavior (eg, frequency, intensity of exercise), and apps that incorporate wearable devices (eg, Fitbit, Garmin) to measure physical activity behavior may be more likely to provide feedback on performance [30]. Therefore, the exclusion of apps that required the use of a wearable device may partially explain the finding of low use of feedback on performance in this study. The most common technical features were *sends reminders* (9/19, 47%) and *allows sharing* (8/19, 42%). There is evidence to suggest that reminders improve the effectiveness of health behavior change interventions [31]; however, other evidence suggests that reminders can hinder habit formation, which may have an impact on long-term behavior change [32]. The high use of sharing is contradictory to the findings that adults find social media features unnecessary and off-putting in health behavior change apps [33]. The removal of sharing features may improve engagement ratings.

The strengths of this study include the systematic search for apps from both App Store and Google Play, the use of an established taxonomy for identifying BCTs, the use of the

MARS instrument to assess the quality of apps, and the inclusion of apps rated above 4.5/5. Further, app ratings were performed by a minimum of two reviewers following best practices for conducting systematic reviews [19]. Good interrater reliability was found for the scoring of BCTs.

The limitations of this study include the exclusion of apps that required use of a wearable device and low interrater reliability for the scoring of app quality through the MARS scale. This may be due to the subjective nature of some sections of the scale. What one reviewer may find aesthetically pleasing, functional, or engaging, another reviewer may not. To account for these differences, the average of the 2 scores was used as the final score for each domain. The temporal relevancy of the results from this study may also be considered a limitation, despite the search being conducted less than 12 months ago because unlike traditional literature where the content remains consistent and unchanged once published, apps are extremely fluid, resulting in frequent updates and modifications to their contents and features. Despite these limitations, this study is the first of its kind and provides valuable real-world findings and implications that are highly relevant to this field as well as the greater audience. Future research should test the overall effectiveness of commercial apps designed to promote physical activity among pregnant women. Finally, research examining the accuracy of app content and the expertise of developers should also be of high priority.

In conclusion, the use of apps for physical activity advice and support in pregnancy is rising. Therefore, an understanding of their quality and inclusion of effective BCTs is required. This is the first study to investigate the quality of popular commercially available apps for promoting physical activity in women during pregnancy. The findings of moderate app quality, with the highest ratings for *functionality* and *aesthetics*, indicate that the apps are user-friendly. However, the low use of evidence-based BCTs for changing physical activity behavior in women during pregnancy indicates that the popular commercial apps currently available may not be effective at promoting physical activity behavior in the pregnant population. More effort needs to be placed on incorporating components most likely to influence behavior change, which is ultimately what they are developed for, while maintaining good functionality. Developers should continue to provide self-monitoring and social comparison and ensure that these components are engaging and effective. In addition, intention formation, goal setting, review of goals, and feedback on performance should be incorporated in new apps and new versions of existing apps.

## Acknowledgments

The authors received no specific funding for this work. Other funding support is as follows: SS is supported by an Early Career Fellowship (GNT1125586) from the National Health and Medical Research Council of Australia and by a Postdoctoral Fellowship (ID 101240) from the National Heart Foundation of Australia. SA is funded by a National Heart Foundation Postdoctoral Fellowship (ID 102609). AA and NC are supported by the National Institutes of Health, which fund the Nutrition Obesity Research Center (P30DK072476) and the National Institute of Nursing Research (R01 NR017644). GPM is supported by a University of Queensland International Postgraduate Research Scholarship. TSN is funded by a Mitacs Fellowship supported by The Society of Obstetricians and Gynecologists of Canada.

LM is funded by a National Heart Foundation Future Leader Fellowship and also funded by the Australian Government's Medical Research Future Fund (MRFF; TABP-18-0001). The MRFF provides funding to support health and medical research and innovation, with the objective of improving the health and well-being of Australians. MRFF funding has been provided to The Australian Prevention Partnership Centre under the MRFF Boosting Preventive Health Research Program. Further information on the MRFF is available in the Australian Government Department of Health website.

---

### Authors' Contributions

MH conceptualized and designed the study, with input from KLA and SS. KLA, SC, and MH screened the apps for eligibility based on app descriptions. All authors downloaded and tested the apps, extracted the data, and contributed to content analysis and quality assessment. KLA and SC identified BCTs and sought confirmation from other reviewers. KLA analyzed the data. KLA, MH, LV, and SA drafted the manuscript. All authors were involved in the interpretation of data and critical revision of the manuscript, and they read and approved the final manuscript.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

Search strategy and search results.

[DOC File, 78 KB - [mhealth\\_v9i4e23649\\_app1.doc](#)]

---

#### Multimedia Appendix 2

App characteristics.

[DOC File, 63 KB - [mhealth\\_v9i4e23649\\_app2.doc](#)]

---

#### Multimedia Appendix 3

Mobile App Rating Scale categories.

[DOC File, 63 KB - [mhealth\\_v9i4e23649\\_app3.doc](#)]

---

#### Multimedia Appendix 4

Mean Mobile App Rating Scale scores and interrater reliability.

[DOC File, 65 KB - [mhealth\\_v9i4e23649\\_app4.doc](#)]

---

#### Multimedia Appendix 5

Behavior change techniques and interrater reliability.

[DOC File, 107 KB - [mhealth\\_v9i4e23649\\_app5.doc](#)]

---

### References

1. Artal R. Exercise during pregnancy and the postpartum period. UpToDate. 2017 Apr. URL: <https://www.uptodate.com/contents/exercise-during-pregnancy-and-the-postpartum-period#> [accessed 2019-01-01]
2. Harrison CL, Brown WJ, Hayman M, Moran LJ, Redman LM. The Role of Physical Activity in Preconception, Pregnancy and Postpartum Health. *Semin Reprod Med* 2016 Mar;34(2):e28-e37 [FREE Full text] [doi: [10.1055/s-0036-1583530](https://doi.org/10.1055/s-0036-1583530)] [Medline: [27169984](https://pubmed.ncbi.nlm.nih.gov/27169984/)]
3. da Silva SG, Ricardo LI, Evenson KR, Hallal PC. Leisure-Time Physical Activity in Pregnancy and Maternal-Child Health: A Systematic Review and Meta-Analysis of Randomized Controlled Trials and Cohort Studies. *Sports Med* 2017 Feb;47(2):295-317. [doi: [10.1007/s40279-016-0565-2](https://doi.org/10.1007/s40279-016-0565-2)] [Medline: [27282925](https://pubmed.ncbi.nlm.nih.gov/27282925/)]
4. 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](https://doi.org/10.1016/j.jsams.2015.09.004)] [Medline: [26481261](https://pubmed.ncbi.nlm.nih.gov/26481261/)]
5. Coll CV, Domingues MR, Gonçalves H, Bertoldi AD. Perceived barriers to leisure-time physical activity during pregnancy: A literature review of quantitative and qualitative evidence. *J Sci Med Sport* 2017 Jan;20(1):17-25. [doi: [10.1016/j.jsams.2016.06.007](https://doi.org/10.1016/j.jsams.2016.06.007)] [Medline: [27372276](https://pubmed.ncbi.nlm.nih.gov/27372276/)]
6. 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](https://doi.org/10.1016/j.wombi.2013.11.001)] [Medline: [24295598](https://pubmed.ncbi.nlm.nih.gov/24295598/)]
7. Lee Y, Moon M. Utilization and Content Evaluation of Mobile Applications for Pregnancy, Birth, and Child Care. *Healthc Inform Res* 2016 Apr;22(2):73-80 [FREE Full text] [doi: [10.4258/hir.2016.22.2.73](https://doi.org/10.4258/hir.2016.22.2.73)] [Medline: [27200216](https://pubmed.ncbi.nlm.nih.gov/27200216/)]
8. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the Next-Generation Perinatal Care: An Overview of the Literature. *J Med Internet Res* 2018 Jun 05;20(6):e202 [FREE Full text] [doi: [10.2196/jmir.9262](https://doi.org/10.2196/jmir.9262)] [Medline: [29871855](https://pubmed.ncbi.nlm.nih.gov/29871855/)]

9. 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](https://doi.org/10.1016/j.wombi.2016.01.008)] [Medline: [26874938](https://pubmed.ncbi.nlm.nih.gov/26874938/)]
10. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. *Int J Behav Nutr Phys Act* 2017 Mar 28;14(1):42 [FREE Full text] [doi: [10.1186/s12966-017-0494-y](https://doi.org/10.1186/s12966-017-0494-y)] [Medline: [28351367](https://pubmed.ncbi.nlm.nih.gov/28351367/)]
11. Michie S, Carey RN, Johnston M, Rothman AJ, de Bruin M, Kelly MP, et al. From Theory-Inspired to Theory-Based Interventions: A Protocol for Developing and Testing a Methodology for Linking Behaviour Change Techniques to Theoretical Mechanisms of Action. *Ann Behav Med* 2018 May 18;52(6):501-512 [FREE Full text] [doi: [10.1007/s12160-016-9816-6](https://doi.org/10.1007/s12160-016-9816-6)] [Medline: [27401001](https://pubmed.ncbi.nlm.nih.gov/27401001/)]
12. 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;8(6):e66385 [FREE Full text] [doi: [10.1371/journal.pone.0066385](https://doi.org/10.1371/journal.pone.0066385)] [Medline: [23799096](https://pubmed.ncbi.nlm.nih.gov/23799096/)]
13. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008 May;27(3):379-387. [doi: [10.1037/0278-6133.27.3.379](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
14. Zinsler LA, Stoll K, Wieber F, Pehlke-Milde J, Gross MM. Changing behaviour in pregnant women: A scoping review. *Midwifery* 2020 Jun;85:102680 [FREE Full text] [doi: [10.1016/j.midw.2020.102680](https://doi.org/10.1016/j.midw.2020.102680)] [Medline: [32151875](https://pubmed.ncbi.nlm.nih.gov/32151875/)]
15. 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](https://doi.org/10.1186/s12966-017-0538-3)] [Medline: [28646889](https://pubmed.ncbi.nlm.nih.gov/28646889/)]
16. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. *Am J Prev Med* 2014 Jun;46(6):649-652. [doi: [10.1016/j.amepre.2014.01.010](https://doi.org/10.1016/j.amepre.2014.01.010)] [Medline: [24842742](https://pubmed.ncbi.nlm.nih.gov/24842742/)]
17. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR Mhealth Uhealth* 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
18. 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](https://doi.org/10.1111/mcn.12768)] [Medline: [30569549](https://pubmed.ncbi.nlm.nih.gov/30569549/)]
19. 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](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
20. 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](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
21. Schoffman DE, Turner-McGrievy G, Jones SJ, Wilcox S. Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? *Transl Behav Med* 2013 Sep;3(3):320-325 [FREE Full text] [doi: [10.1007/s13142-013-0206-3](https://doi.org/10.1007/s13142-013-0206-3)] [Medline: [24073184](https://pubmed.ncbi.nlm.nih.gov/24073184/)]
22. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *Int J Behav Nutr Phys Act* 2016 Mar 10;13:35 [FREE Full text] [doi: [10.1186/s12966-016-0359-9](https://doi.org/10.1186/s12966-016-0359-9)] [Medline: [26964880](https://pubmed.ncbi.nlm.nih.gov/26964880/)]
23. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC Public Health* 2014 Jun 25;14:646 [FREE Full text] [doi: [10.1186/1471-2458-14-646](https://doi.org/10.1186/1471-2458-14-646)] [Medline: [24965805](https://pubmed.ncbi.nlm.nih.gov/24965805/)]
24. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. *Int J Behav Nutr Phys Act* 2014 Jul 25;11:97 [FREE Full text] [doi: [10.1186/s12966-014-0097-9](https://doi.org/10.1186/s12966-014-0097-9)] [Medline: [25059981](https://pubmed.ncbi.nlm.nih.gov/25059981/)]
25. Hayes AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures* 2007 Apr;1(1):77-89. [doi: [10.1080/19312450709336664](https://doi.org/10.1080/19312450709336664)]
26. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Diequez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform* 2016 Apr;60:243-251 [FREE Full text] [doi: [10.1016/j.jbi.2016.02.002](https://doi.org/10.1016/j.jbi.2016.02.002)] [Medline: [26903153](https://pubmed.ncbi.nlm.nih.gov/26903153/)]
27. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
28. 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](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
29. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010 Feb 17;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]



30. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res* 2014 Aug 15;16(8):e192 [FREE Full text] [doi: [10.2196/jmir.3469](https://doi.org/10.2196/jmir.3469)] [Medline: [25131661](https://pubmed.ncbi.nlm.nih.gov/25131661/)]
31. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. *J Med Internet Res* 2009 May 14;11(2):e16 [FREE Full text] [doi: [10.2196/jmir.1138](https://doi.org/10.2196/jmir.1138)] [Medline: [19632970](https://pubmed.ncbi.nlm.nih.gov/19632970/)]
32. Stawarz K, Cox A, Blandford A. Beyond Self-Tracking and Reminders: Designing Smartphone Apps That Support Habit Formation. 2015 Presented at: 33rd Annual AMC Conference on Human Factors in Computing Systems; April, 2015; Seoul, Republic of Korea. [doi: [10.1145/2702123.2702230](https://doi.org/10.1145/2702123.2702230)]
33. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]

## Abbreviations

**BCT:** behavior change technique

**K $\alpha$ :** Krippendorff's alpha

**MARS:** Mobile App Rating Scale

*Edited by G Eysenbach; submitted 18.08.20; peer-reviewed by B Wollesen, P Grace-Farfaglia; comments to author 01.10.20; revised version received 21.10.20; accepted 17.01.21; published 07.04.21.*

*Please cite as:*

Hayman M, Alfrey KL, Cannon S, Alley S, Rebar AL, Williams S, Short CE, Altazan A, Comardelle N, Currie S, Denton C, Harrison CL, Lamerton T, Mena GP, Moran L, Mottola M, Nagpal TS, Vincze L, Schoeppe S

*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;9(4):e23649

URL: <https://mhealth.jmir.org/2021/4/e23649>

doi: [10.2196/23649](https://doi.org/10.2196/23649)

PMID: [33825693](https://pubmed.ncbi.nlm.nih.gov/33825693/)

©Melanie Hayman, Kristie-Lee Alfrey, Summer Cannon, Stephanie Alley, Amanda L Rebar, Susan Williams, Camille E Short, Abby Altazan, Natalie Comardelle, Sinead Currie, Caitlin Denton, Cheryce L Harrison, Tayla Lamerton, Gabriela P Mena, Lisa Moran, Michelle Mottola, Taniya S Nagpal, Lisa Vincze, Stephanie Schoeppe. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 07.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Usability and Acceptability of a Mobile App for the Self-Management of Alcohol Misuse Among Veterans (Step Away): Pilot Cohort Study

Carol A Malte<sup>1,2</sup>, MSW; Patrick L Dulin<sup>3</sup>, PhD; John S Baer<sup>1,2,4</sup>, PhD; John C Fortney<sup>2,5</sup>, PhD; Anissa N Danner<sup>1,2</sup>, MSW; Aline M K Lott<sup>1,2</sup>, MA; Eric J Hawkins<sup>1,2,5</sup>, PhD

<sup>1</sup>Center of Excellence in Substance Addiction Treatment and Education (CESATE), Veterans Affairs Puget Sound Health Care System, Seattle, WA, United States

<sup>2</sup>Health Services Research & Development Seattle Center of Innovation for Veteran-Centered and Value-Driven Care, Veterans Affairs Puget Sound Health Care System, Seattle, WA, United States

<sup>3</sup>Department of Psychology, University of Alaska Anchorage, Anchorage, AK, United States

<sup>4</sup>Department of Psychology, University of Washington, Seattle, WA, United States

<sup>5</sup>Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

**Corresponding Author:**

Carol A Malte, MSW

Center of Excellence in Substance Addiction Treatment and Education (CESATE)

Veterans Affairs Puget Sound Health Care System

1660 S Columbian Way, MS 116 ATC

Seattle, WA, 98108

United States

Phone: 1 206 277 3780

Email: [carol.malte@va.gov](mailto:carol.malte@va.gov)

## Abstract

**Background:** Alcohol misuse is common among Operation Enduring Freedom and Operation Iraqi Freedom veterans, yet barriers limit treatment participation. Mobile apps hold promise as means to deliver alcohol interventions to veterans who prefer to remain anonymous, have little time for conventional treatments, or live too far away to attend treatment in person.

**Objective:** This pilot study evaluated the usability and acceptability of Step Away, a mobile app designed to reduce alcohol-related risks, and explored pre-post changes on alcohol use, psychological distress, and quality of life.

**Methods:** This single-arm pilot study recruited Operation Enduring Freedom and Operation Iraqi Freedom veterans aged 18 to 55 years who exceeded National Institute on Alcohol Abuse and Alcoholism drinking guidelines and owned an iPhone. Enrolled veterans (N=55) completed baseline and 1-, 3-, and 6-month assessments. The System Usability Scale (scaled 1-100,  $\geq 70$  indicating acceptable usability) assessed the effectiveness, efficiency, and satisfaction dimensions of usability, while a single item (scaled 1-9) measured the attractiveness of 10 screenshots. Learnability was assessed by app use during week 1. App engagement (proportion of participants using Step Away, episodes of use, and minutes per episode per week) over 6 months measured acceptability. Secondary outcomes included pre-post change on heavy drinking days (men:  $\geq 5$  drinks per day; women:  $\geq 4$  drinks per day) and Short Inventory of Problems–Revised, Kessler-10, and brief World Health Organization Quality of Life Questionnaire scores.

**Results:** Among the 55 veterans enrolled in the study, the mean age was 37.4 (SD 7.6), 16% (9/55) were women, 82% (45/55) were White, and 82% (45/55) had an alcohol use disorder. Step Away was used by 96% (53/55) of participants in week 1, 55% (30/55) in week 4, and 36% (20/55) in week 24. Step Away use averaged 55.1 minutes (SD 57.6) in week 1 and <15 minutes per week in weeks 2 through 24. Mean System Usability Scale scores were 69.3 (SD 19.7) and 71.9 (SD 15.8) at 1 and 3 months, respectively. Median attractiveness scores ranged from 5 to 8, with lower ratings for text-laden screens. Heavy drinking days decreased from 29.4% (95% CI 23.4%-35.4%) at baseline to 16.2% (95% CI 9.9%-22.4%) at 6 months ( $P < .001$ ). Likewise, over 6 months, Short Inventory of Problems–Revised scores decreased from 6.3 (95% CI 5.1-7.5) to 3.6 (95% CI 2.4-4.9) ( $P < .001$ ) and Kessler-10 scores decreased from 18.8 (95% CI 17.4-20.1) to 17.3 (95% CI 15.8-18.7) ( $P = .046$ ). Changes were not detected on quality of life scores.

**Conclusions:** Operation Enduring Freedom and Operation Iraqi Freedom veterans found the usability of Step Away to be acceptable and engaged in the app over the 6-month study. Reductions were seen in heavy drinking days, alcohol-related problems, and Kessler-10 scores. A larger randomized trial is warranted to confirm our findings.

(*JMIR Mhealth Uhealth* 2021;9(4):e25927) doi:[10.2196/25927](https://doi.org/10.2196/25927)

## KEYWORDS

mobile apps; alcohol misuse; smartphone; veterans; access

## Introduction

Alcohol misuse, which is associated with a number of adverse social, economic, and health-related consequences [1,2], is one of the most common conditions among Operation Enduring Freedom (OEF) and Operation Iraqi Freedom (OIF) service members and veterans, with estimates ranging from 22% to 40% [3-5]. Previous reports suggest the rate of alcohol misuse among OEF and OIF veterans is 2 times the rate observed for similarly aged veterans who did not serve in OEF and OIF and much higher among men (22%) than women (5%) [5]. Despite high rates of use, many OEF and OIF veterans with alcohol-related problems do not receive alcohol-related care or only receive care after significant delay [6-8]. The large gap between those who need and those who receive treatment is thought to be due in part to barriers to using available services [9].

OEF and OIF veterans have reported several impediments to seeking mental health care, including stigma-related barriers, such as beliefs that they will be perceived as weak for seeking help, and logistical barriers, such as insufficient time, burdensome paperwork, and long distances to the nearest treatment facility [3,10,11]. Mobile apps, delivered on smartphones, have the potential to address the majority of these barriers, as they can deliver alcohol interventions to those who are interested in help but prefer to remain anonymous, have little time for traditional therapy, or must travel too far to attend treatment in person [12]. Unlike traditional alcohol interventions that rely on patient-to-provider encounters, mobile apps deliver timely interventions to individuals in their natural settings at crucial moments when the need for intervention is high and repeatedly over time [13]. Further, smartphone ownership in the United States is common, particularly among persons aged 18 to 49 years [14], and there is considerable interest in tools that allow for self-management of alcohol use [15].

While mobile apps to manage alcohol use have proliferated in recent years, they vary greatly in quality, and many are narrow in scope (ie, limited to tracking of alcohol consumed, estimating blood alcohol content [15]) or lack key evidence-based intervention components associated with behavior change (ie, normative feedback, use of social support [16]). While data on the effect of apps, particularly those aimed at young adults, on drinking outcomes are growing [17,18], limited information exists on their acceptability and usability, particularly among OEF and OIF veterans [19,20]. The importance of collecting this type of information is highlighted by a recent qualitative study of rural veterans [21], which found that these veterans held more negative views of apps relative to urban veterans and expressed that apps were hard to navigate, hard to access due

to connectivity issues, and opposed to their values (impersonal, lacking connection to community). App acceptability and usability are closely related to engagement, a key issue in app development, with several studies finding that app participation drops off steeply after 1 week, with only the most at-risk and motivated individuals remaining engaged (eg, Drinkaware [22]). The relationship between app engagement and drinking outcomes is not well established. A small study of an app (InDEx) developed for British ex-service members showed promise in engaging participants over the 4-week study course, with modest reductions in drinking [23]. A second study found that age, gender, and education rather than baseline drinking levels were associated with app (Drink Less) engagement in a nonveteran sample; however, engagement with the app was not associated with drinking outcomes [24].

Two of the most studied and comprehensive smartphone-based interventions are the Location-Based Monitoring and Intervention System for Alcohol Use Disorders (LBMI-A) and the Addiction Comprehensive Health Enhancement Support System (A-CHESS). Designed to provide stand-alone treatment, LBMI-A has been evaluated among a small community sample who met Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision (DSM-5) criteria for an alcohol use disorder (AUD). Results suggest a significant reduction in the percentage of heavy drinking days (HDD) from 56% at baseline to 25% at 6 weeks [25,26]. A substantial barrier to engagement with LBMI-A was the reported difficulty in simultaneously using a personal and LBMI-A-enabled phone in the pilot study (prior to mobile apps' emergence). In contrast to LBMI-A, A-CHESS was designed to prevent relapse to heavy drinking among patients who completed residential treatment for an AUD [27]. A-CHESS patient-generated data are accessed by treatment providers to monitor patient progress; as such, it is not anonymous or independent of traditional treatment.

Step Away, the next iteration of LBMI-A, is an iOS-based mobile app designed to help people self-manage drinking and alcohol-related problems and thus is an alternative for those who do not want or have access to traditional treatment [12]. A recent systematic review of eHealth interventions for alcohol identified several recommendations [20], such as addressing the spectrum of alcohol use (at-risk drinking to AUD), including cognitive behavioral coping strategies and access to timely interventions for skill development and navigation of high-risk situations, and using mobile technologies such as push notifications to promote engagement. Step Away complies with many of these recommendations by addressing a range of severity of alcohol-related problems; providing education, advice, and goal setting related to alcohol use; soliciting participant data on drinking, craving, and mood through push

notifications; and compiling these data in regular feedback reports.

This mixed methods pilot study was designed to evaluate the usability and acceptability of Step Away among OEF and OIF veterans with alcohol misuse who are not involved in traditional alcohol treatment. The primary aim of this project was to assess the usability of Step Away among OEF and OIF veterans. According to the International Organization for Standardization Guidance on Usability [28], key dimensions of usability are efficiency, effectiveness, and satisfaction. Further, a recent review of mobile health (mHealth) apps [29] recommends the assessment of attractiveness and learnability, important usability dimensions that are infrequently evaluated in the literature. A second aim was to assess the acceptability of the app, which was measured by engagement. Exploratory aims were to evaluate change in alcohol use, psychological distress, and health-related quality of life outcomes associated with the use of Step Away over the 6-month study course and to examine associations between app engagement, baseline characteristics, and drinking outcomes.

## Methods

This single-arm prospective cohort study involving a national sample of OEF and OIF veterans ( $n=55$ ) occurred between September 2017 and December 2018. The study was approved by the Veteran Affairs (VA) Puget Sound Institutional Review Board.

### Study Participants

OEF and OIF veterans were eligible to participate if they (1) were aged 18 to 55 years, (2) screened positive for alcohol misuse ( $\geq 5$  standard drinks on any day or  $\geq 15$  standard drinks per week for men and  $\geq 4$  standard drinks on any day or  $\geq 8$  standard drinks per week for women) in the prior 4 weeks, and (3) owned an iPhone. Exclusion criteria included (1) participation in substance use treatment, including Alcoholics Anonymous, in the past 30 days; (2) pregnancy; and (3) satisfaction of DSM-5 criteria for a drug use disorder (nicotine use disorders and drug disorders in remission allowed) or psychotic disorder.

### Procedures

The planned strategy of recruiting participants using social media advertisements proved unsuccessful, with advertisements yielding no contacts from interested participants after 3 months. Subsequently, investigators used the VA Corporate Data Warehouse, a data repository of patient-level data from the VA electronic medical record, to identify OEF and OIF veterans with Alcohol Use Disorders Identification Test–Consumption [30] scores of 5 or greater (indicating alcohol misuse) who used VA care in the past 6 months ( $n=13,266$ ) for recruitment by letters and follow-up contacts. Potentially eligible veterans were randomly selected and, in blocks of 100, mailed recruitment letters introducing the study as an evaluation of an app designed to improve the health of individuals who drink alcohol. A total of 10% of letters were sent to women to ensure adequate sampling. Research staff then attempted to contact veterans weekly for up to 3 weeks via telephone and conducted brief

telephone screens to assess eligibility with those who were reached and interested in participation. Eligible and interested veterans were mailed a consent form and completed informed consent procedures by telephone. Following consent, veterans completed a baseline research assessment and then received instructions on how to download Step Away to their smartphone. Follow-up assessments occurred at 1, 3, and 6 months. A subset of veterans completed qualitative interviews during the initial 3 months (data to be reported separately). All assessments were completed by telephone, with secure messaging (SMS) used to schedule and remind participants about appointments. Participants were compensated US \$40 for completing baseline and 3- and 6-month follow-up assessments and US \$5 for completing a usability questionnaire at the 1- and 3-month follow-up.

### Intervention: Step Away Mobile App Content and Features

A version of the Step Away app similar to the one used by this study has been described in detail elsewhere [31]. Briefly, Step Away is a comprehensive intervention for alcohol misuse based on the evidence-based principles of motivational enhancement therapy (MET) [32], relapse prevention [33], and community reinforcement [34,35]. Consistent with a harm reduction approach that supports motivation to change in a nonjudgmental, facilitative manner, Step Away supports both reductions in and abstinence from drinking. Step Away incorporates relapse prevention strategies to identify and cope with situations that increase the risk of relapsing or drinking inconsistently with goals as well as community reinforcement strategies that encourage involvement of supportive others in setting drinking goals and participating in nondrinking activities.

In total, Step Away consists of 10 modules: drinking profile, goal setting, rewards, cravings, strategies, supportive others, reminders, high-risk times, moods, and new activities. Taken together, these modules encompass assessment, normative feedback, goal setting, behavioral strategies tailored to the goals of moderation or abstinence, involvement of supportive others, and replacement activities. Users can personalize the app through the use of reminders about reasons for changes, high-risk times, and scheduled activities. It is anticipated that app users will spend more time in the app during the initial session to review content and set up goals and personal preferences.

Using push notifications, Step Away prompts users daily to complete a brief self-monitoring questionnaire on current mood, drinking behaviors, and alcohol cravings and triggers for alcohol use during the prior 24 hours. Step Away provides weekly feedback, highlighting progress toward goals, ongoing cravings and triggers, and moods experienced throughout the prior week. Feedback includes strategies for managing common alcohol cravings and triggers and a feature for users to enter future high-risk events that might negatively influence drinking goals. Step Away provides real-time intervention options, including strategies for managing cravings or negative emotions, requests for help (eg, to manage anxiety or alcohol cravings), and working through a problem via a problem-solving algorithm [36]. Veterans are provided with an option to contact the

Veterans Crisis Line if in crisis and VA or non-VA addiction programs if more traditional treatment services are desired.

## Measures and Outcomes

Study data included both patient self-report and data generated automatically through the use of Step Away. Assessments are detailed below. Step Away use data were automatically generated and stored on secure Step Away servers. Data included date and time of session start and end, specific screen views, time spent per screen view, use of Step Away features, and completion of daily brief assessments.

## Eligibility and Patient Characteristics

The brief telephone screen assessed age, iPhone ownership, OEF or OIF service, pregnancy, current participation in VA or non-VA addiction treatment, recent alcohol and drug use, and potential psychotic symptoms. Positive screens for drug use or psychotic symptoms were followed up with the appropriate modules of the Mini International Neuropsychiatric Interview (MINI) modified for the DSM-5 [37]. At baseline, demographics and the MINI depressive, bipolar, anxiety, alcohol use, and posttraumatic stress disorders modules were completed.

## Primary Outcomes

The main study outcome was app usability, comprised of effectiveness, efficiency, satisfaction, learnability, and attractiveness. Effectiveness measures the ability of users to perform a given task, efficiency describes the resources (eg, time) expended to perform the task after it has been learned, and satisfaction represents users' assessments of how well the device met their needs. These 3 dimensions of usability were assessed at the 1- and 3-month follow-up using the System Usability Scale (SUS; scaled 1 to 100), a 10-item, well-validated questionnaire [38] with scores of  $\geq 70$  considered acceptable [39]. Learnability represents users' abilities to accomplish a task on their first attempt, which is an important determinant of engagement. The learnability of Step Away was measured by the total time spent in the app on the first episode of use and the first week of use. Further, an analysis of the SUS validated a 2-factor structure capturing learnability and usability [40]. Similar to the full SUS, both components are scaled from 0 to 100. Additionally, the attractiveness of Step Away was assessed at 1 month using a validated 9-point Likert scale item (scaled 1=low to 9=high) designed to assess the visual appeal of web pages [41]. Median scores on the 1-item questionnaire measured attractiveness of 10 distinct Step Away screenshots.

Acceptability was measured through app engagement [42] and included the following measures derived from Step Away use data: proportion of participants who used Step Away, number of episodes of use, and episode length (time between opening and closing of app) overall and by week during the 6 months after baseline assessment. We also calculated the number of times each module was viewed and the number of times daily questionnaires were completed by each participant. Means (standard deviations) and medians (interquartile ranges) were used to calculate these measures.

## Secondary Outcomes

Secondary outcomes included change in alcohol use, psychological distress, and health-related quality of life from baseline to month 6. Assessments collected at baseline and 3- and 6-month follow-up included (1) the Timeline Followback (TLFB) [43] to assess alcohol use in the prior 30 days and number of standard drinks (14 grams of alcohol) per day; (2) the Short Inventory of Problems–Revised (SIP-R) [44] to assess drinking-related consequences; (3) the Kessler-10, a reliable and valid 10-item psychological distress measure (with scores ranging from 10 to 50), to assess distress over the prior month [45,46]; and (4) the abbreviated World Health Organization Quality of Life (WHOQOL-BREF), a 26-item measure derived from the WHOQOL-100 quality of life measure, to assess 4 domains related to quality of life, that is, physical health, psychological health, social relationships, and environment [47].

Alcohol use outcomes were defined as changes in the percentage of HDD ( $\geq 5$  drinks per day for men;  $\geq 4$  drinks per day for women) in the prior 30 days and the proportion of patients drinking below the recommended limits, defined as  $< 15$  drinks per week and  $< 5$  drinks per day for men and  $< 8$  drinks per week and  $< 4$  drinks per day for women, as determined by TLFB data. Change in heavy drinking days was selected because it is a marker of alcohol misuse, associated with long-term health outcomes, and frequently used as an outcome in the alcohol treatment literature [48,49]. Alcohol consequences were measured by change in the SIP-R total score. Changes in the Kessler-10 total score and 4 domain scores of the WHOQOL-BREF measured psychological distress and health-related quality of life outcomes, respectively.

## Analyses

Usability, learnability, attractiveness, and engagement outcomes as well as participant characteristics are presented using descriptive statistics, including means with standard deviations and medians with interquartile ranges. Associations between app engagement (total weeks and minutes of use) and demographics, baseline alcohol use, drug use, and mental health diagnoses were examined using unadjusted linear regression. Responses at baseline and 3 and 6 months represent repeated measures of clinical outcomes over time. Longitudinal analyses using multilevel mixed models were used to estimate the percent change in HDD between baseline and month 6, with a fixed effect for time and a unique patient identifier included as a random effect. Gender, age, and binary measures of AUD, any drug use, and mental health disorders as per the MINI (major depression, generalized anxiety, and posttraumatic stress disorder) at baseline were included as covariates to determine their associations with changes in the percentage of HDD. A similar approach was used to estimate changes in alcohol consequences (SIP-R), distress (Kessler-10), and quality of life (WHOQOL-BREF) outcomes. The proportion of participants drinking below guideline limits at follow-up was examined using a Pearson chi-square test. Associations between change over time in alcohol-related outcomes (HDD, SIP-R) and app engagement were explored using multilevel mixed models that included time, a measure of app engagement (total weeks or minutes of use), and the interaction between time and app

engagement, with a patient identifier included as a random effect. All analyses were conducted with Stata MP (version 15; StataCorp).

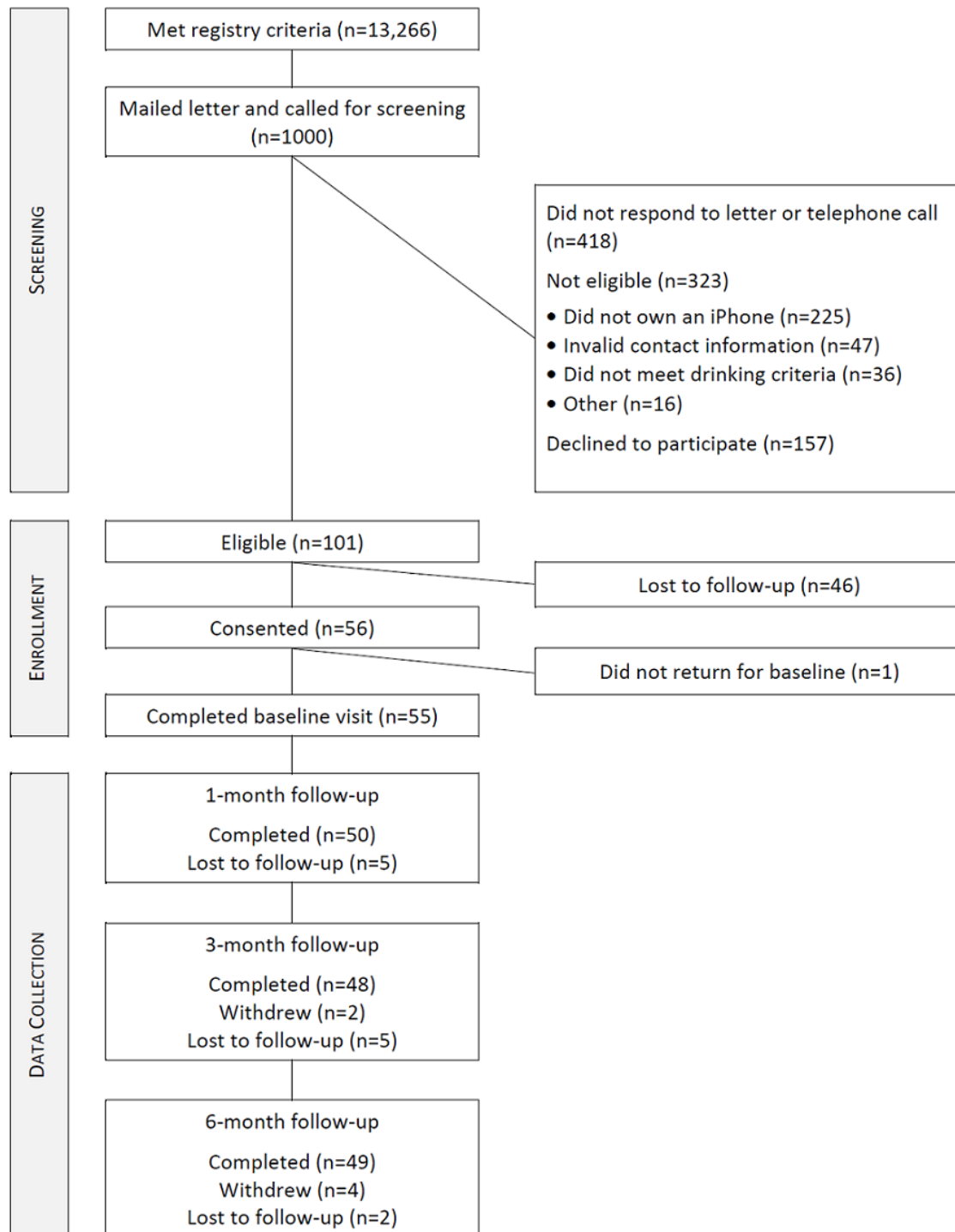
## Results

### Participants

As shown in [Figure 1](#), over 140 days, 1000 potentially eligible veterans were invited to participate by mail. Of this group, 621

(62.1%) veterans did not respond or were unable to be contacted by telephone (418/1000, 41.8%) or declined to participate (157/1000, 15.7%), and 324 (32.4%) screened ineligible (eg, did not own an iPhone). Of the remaining 101 veterans, 55 enrolled in the study. Completion rates of follow-up assessments were over 85% at all time points. Participants' mean age was 37.4 (SD 7.6), 16% (9/55) were women, and 82% (45/55) were White ([Table 1](#)). A total of 82% (45/55) of participants met DSM-5 criteria for an alcohol use disorder and 33% (18/55) had received treatment for a substance use disorder in the past.

**Figure 1.** CONSORT diagram.



**Table 1.** Participant characteristics (N=55).

Characteristic	Value
Age, mean (SD)	37.4 (7.6)
Female, n (%)	9 (16.4)
<b>Race, n (%)</b>	
White	45 (81.8)
Black	4 (7.3)
Asian or Pacific Islander	3 (5.5)
American Indian	3 (5.5)
Hispanic <sup>a</sup>	10 (18.5)
<b>Education (years), mean (SD)</b>	
12	11 (20.0)
13-15	18 (32.7)
16+	26 (47.3)
<b>Marital status, n (%)</b>	
Married or living with partner	36 (65.4)
Divorced or separated	8 (14.6)
Never married	11 (20.0)
Currently employed, n (%)	43 (78.2)
<b>Income (US \$), n (%)</b>	
0 to 10,000	2 (3.6)
10,000 to 50,000	25 (45.5)
50,000 to 100,000	21 (38.2)
>100,000	7 (12.7)
Prior substance use disorder treatment, n (%)	18 (32.7)
Any drug use, n (%)	2 (3.6)
<b>Mental health conditions, n (%)</b>	
Current major depressive disorder	4 (7.3)
Generalized anxiety disorder	3 (5.5)
Posttraumatic stress disorder	5 (9.1)
Alcohol use disorder	45 (81.8)

<sup>a</sup>N=54.

### Usability of Step Away

Participants' mean SUS scores, which assess effectiveness, efficiency, and satisfaction, were 69.3 (SD 19.7) and 71.9 (SD 15.8) at the 1- and 3-month follow-up, respectively, with 62% (31/50) and 77% (36/47) of participants scoring  $\geq 70$  points at the 2 time points. Learnability was assessed by time taken to complete modules the first time. Participants spent an average 19.2 (SD 25.1) minutes in Step Away for the first episode of use compared to an average of 11.9 (SD 10.6) minutes per episode during all of week 1 and less than a minute per episode at week 24. Mean scores on the learnability component of the SUS were 80.0 (SD 22.9) at 1 month and 79.0 (SD 17.3) at 3 months. Of the 10 screenshots that participants evaluated,

attractiveness scores ranged from a median of 5 to 8 (out of 9), with lower ratings given to text-laden screens.

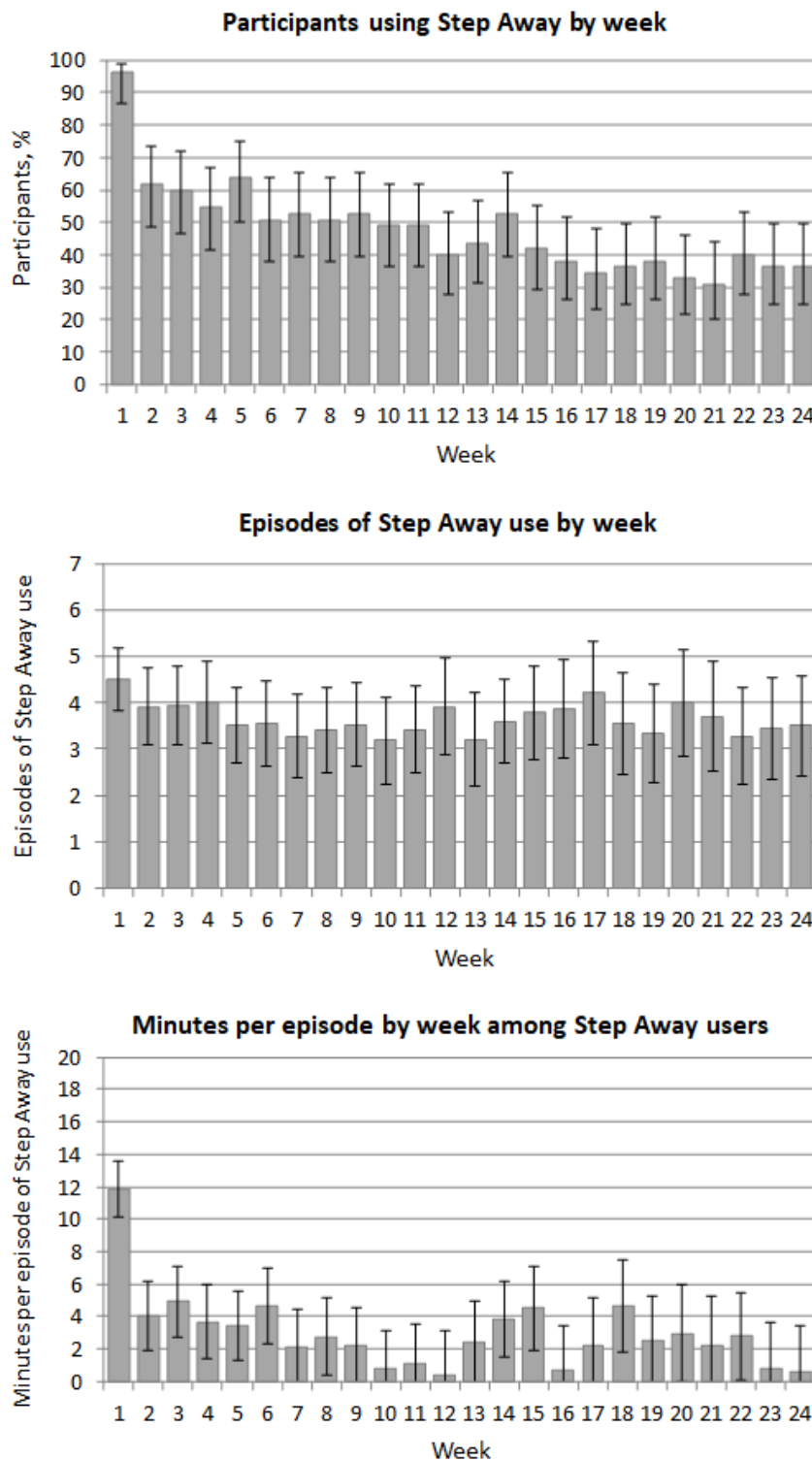
### Acceptability of Step Away

As seen in Figure 2, nearly all participants accessed Step Away in week 1, with the percentage dropping to 62% (34/55) in week 2 and 36% (20/55) by week 24. All 55 participants accessed the app at least once over the 6-month course of the study. A total of 24 of the 55 (44%) participants accessed the app every week during weeks 1 through 4, and 12 (22%) accessed the app every week during weeks 1 through 12. Time spent in the app dropped off considerably after week 1, with participants spending close to 60 minutes (mean 55.1, SD 57.6) in Step Away in week 1 and less than 15 minutes in weeks 2 to 24. However, episodes of Step Away use per week remained at 3 to 4 over the course

of the study among participants accessing the app at least once in the given week. The majority of app modules were accessed by  $\geq 60\%$  of participants, and consumption- and tracking-focused

modules such as daily feedback were accessed most frequently (median times accessed: 12, IQR 2-52) and by the largest proportion of participants ( $>80\%$ ) (Table 2).

**Figure 2.** Step Away app engagement by week. App users are defined as study participants who used the app in a given week.





**Table 2.** Participants' use of Step Away by module from baseline to month 6.

Feature and activity	Ever accessed, n (%)	Number of times accessed	
		Mean (SD)	Median (IQR)
<b>General overview</b>			
Overview and setup	55 (100)	2.3 (1.5)	2 (1-3)
<b>Drinking profile module</b>			
Initial drinking profile	53 (96.4)	2.9 (2.0)	2 (2-4)
Maximum drinks	53 (96.4)	1.6 (1.2)	1 (1-2)
Peak BAC <sup>a</sup>	53 (96.4)	1.4 (0.7)	1 (1-2)
Dependency and SADQ <sup>b</sup>	49 (89.1)	0.9 (0.4)	1 (1-1)
Money and total cost	47 (85.5)	0.9 (0.4)	1 (1-1)
<b>Goals module</b>			
Goals step	43 (78.2)	2.2 (2.2)	2 (1-3)
Goals path (abstinence and moderation)	42 (76.4)	0.9 (0.7)	1 (1-1)
<b>Rewards module</b>			
Rewards	41 (74.5)	1.3 (1.7)	1 (0-2)
<b>Craving module</b>			
Craving	35 (63.6)	1.0 (1.0)	1 (0-2)
<b>Supportive others module</b>			
Supportive person	34 (61.8)	1.4 (1.7)	1 (0-2)
<b>Strategies module</b>			
Strategies	33 (60.0)	1.1 (1.3)	1 (0-2)
<b>Reminders module</b>			
Reminder	33 (60.0)	0.8 (0.9)	1 (0-1)
<b>Moods module</b>			
Mood	34 (61.8)	1.0 (1.3)	1 (0-1)
Depression level	33 (60.0)	0.6 (0.6)	1 (0-1)
Stress level	33 (60.0)	0.6 (0.6)	1 (0-1)
<b>High-risk times module</b>			
High-risk accessed	33 (60.0)	0.9 (1.1)	1 (0-1)
<b>New activities module</b>			
Activities step accessed	33 (60.0)	1.0 (1.1)	1 (0-2)
Activities entered	32 (58.2)	0.8 (0.8)	1 (0-1)
<b>Ongoing assessment and feedback</b>			
Daily feedback finished	45 (81.8)	36.9 (51.7)	12 (2-52)
Weekly feedback finished	35 (63.6)	7.3 (9.1)	3 (0-14)
Weekly feedback drink total	34 (61.8)	5.8 (7.8)	2 (0-11)
<b>In-the-moment tools</b>			
Get help accessed	46 (83.6)	2.6 (2.5)	2 (1-3)
Help type selected (anxious, craving, down, problem, other call)	15 (27.3)	0.4 (0.8)	0 (0-1)

<sup>a</sup>BAC: blood alcohol concentration.

<sup>b</sup>Severity of Alcohol Dependence Questionnaire.

Associations with app engagement, as measured by total weeks and minutes of app use over 24 weeks, are shown in Table 3. Total weeks of use was positively associated with female gender, whereas total minutes of app use was positively associated with

age, years of education, and income. Associations were not detected between engagement and baseline alcohol use severity, drug use, or mental health comorbidity.

**Table 3.** Associations between participant characteristics and Step Away engagement.

Characteristic	Weeks of use		Total minutes of use	
	$\beta$ (95% CI)	<i>P</i> value	$\beta$ (95% CI)	<i>P</i> value
Age	0.16 (–0.12 to 0.45)	.25	6.46 (1.44 to 11.47)	.01
Female	7.58 (2.42 to 12.74)	.005	50.79 (–63.66 to 165.24)	.38
Person of color	1.00 (–3.72 to 5.73)	.67	30.54 (–56.47 to 117.55)	.49
<b>Education</b>				
12 years	Ref <sup>a</sup>	N/A <sup>b</sup>	Ref	N/A
13-15 years	3.94 (–1.23 to 9.11)	.13	56.02 (–14.31 to 126.35)	.12
16+ years	5.12 (–0.08 to 10.32)	.05	109.89 (40.28 to 179.49)	.003
Income >\$50,000 (US)	1.00 (–3.55 to 5.55)	.66	88.20 (14.93 to 161.48)	.02
<b>Baseline characteristics</b>				
Count AUD <sup>c</sup> criteria	–0.61 (–1.33 to 0.10)	.09	–12.07 (–25.74 to 1.59)	.08
Percent heavy drinking days	–0.01 (–0.08 to 0.06)	.76	–0.26 (–1.41 to 0.89)	.65
Average drinks per day	–0.35 (–1.18 to 0.48)	.41	–3.39 (–16.08 to 9.30)	.59
SIP-R <sup>d</sup> score	–0.10 (–0.49 to 0.29)	.62	–2.33 (–8.42 to 3.76)	.45
Drug use	–2.01 (–13.09 to 9.07)	.72	–71.88 (–152.04 to 8.28)	.08
Mental health condition	–0.04 (–6.03 to 5.94)	.99	57.50 (–66.80 to 181.81)	.36

<sup>a</sup>Ref: reference category.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>AUD: alcohol use disorder.

<sup>d</sup>SIP-R: Short Inventory of Problems–Revised.

### Alcohol Use, Psychological Distress, and Health-Related Quality of Life

The percentage of participants drinking above recommended limits dropped from 100% (55/55) at baseline to 88% (42/48) at 3 months and 80% (39/49) at 6 months ( $P=.003$ ). Participants' scores over time are presented in Table 4 and visually displayed in Figure 3. Following adjustment for covariates, decreases were seen in participants' percentage of HDD in the past 30 days from 29.4% at baseline to 21.8% at 3 months ( $\beta=-7.6$ , 95% CI –13.6 to –1.6;  $P=.01$ ) and 16.2% at 6 months ( $\beta=-13.2$ , 95% CI –19.2 to –7.2;  $P<.001$ ). Likewise, drinking

consequences as per the SIP-R decreased from 6.3 at baseline to 4.4 at 3 months ( $\beta=-1.9$ , 95% CI –2.8 to –0.9;  $P<.001$ ) and 3.6 at 6 months ( $\beta=-2.7$ , 95% CI –3.6 to –1.7;  $P<.001$ ). Psychological distress as per the Kessler-10 scores decreased from baseline an estimated –1.9 (95% CI –3.3 to –0.4;  $P=.01$ ) points at 3 months and –1.5 (95% CI –2.9 to 0.0;  $P=.046$ ) points at 6 months. Changes were not detected in any of the WHOQOL-BREF domain scores over the study course. No associations were detected between changes in alcohol-related outcomes and app engagement (total weeks and minutes of app use over 24 weeks).

**Table 4.** Estimated alcohol use, psychological distress, and health-related quality of life scores over time.<sup>a</sup>

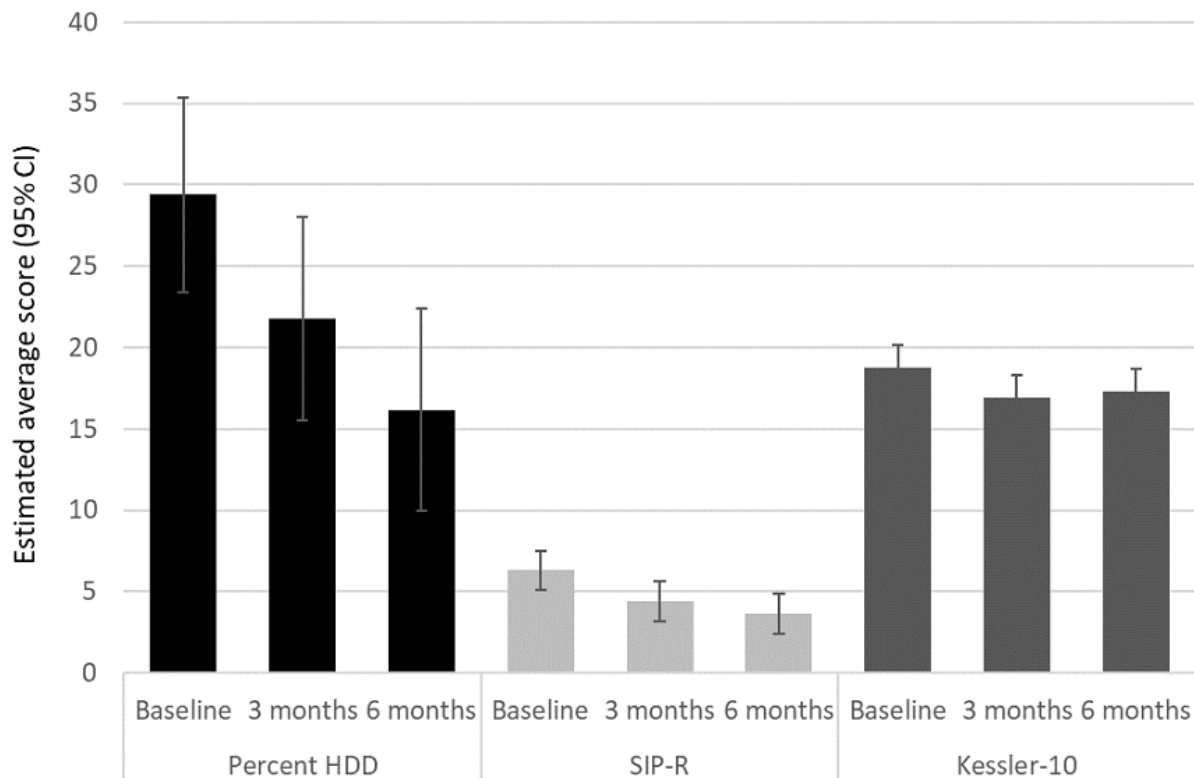
Measure	Estimated mean (95% CI)
<b>Heavy drinking days, %</b>	
Baseline	29.4 (23.4-35.4)
Month 3	21.8 (15.5-28.0)
Month 6	16.2 (9.9-22.4)
<b>SIP-R<sup>b</sup> score</b>	
Baseline	6.3 (5.1-7.5)
Month 3	4.4 (3.2-5.7)
Month 6	3.6 (2.4-4.9)
<b>Kessler-10</b>	
Baseline	18.8 (17.4-20.1)
Month 3	16.9 (15.4-18.3)
Month 6	17.3 (15.8-18.7)
<b>WHOQOL-BREF<sup>c</sup>, physical health</b>	
Baseline	59.5 (57.1-61.9)
Month 3	59.1 (56.5-61.6)
Month 6	60.2 (57.7-62.8)
<b>WHOQOL-BREF, mental health</b>	
Baseline	67.9 (64.8-71.0)
Month 3	67.5 (64.3-70.7)
Month 6	68.0 (64.8-71.2)
<b>WHOQOL-BREF, social relationships</b>	
Baseline	72.6 (67.4-77.8)
Month 3	75.4 (70.0-80.8)
Month 6	73.7 (68.3-79.0)
<b>WHOQOL-BREF, environment</b>	
Baseline	79.9 (76.7-83.1)
Month 3	78.2 (74.9-81.5)
Month 6	80.1 (76.7-83.4)

<sup>a</sup>Adjusted for gender, age, and binary measures of alcohol use disorder, any drug use, and mental health disorders (major depression, generalized anxiety, and posttraumatic stress disorder) at baseline.

<sup>b</sup>SIP-R: Short Inventory of Problems–Revised.

<sup>c</sup>WHOQOL-BREF: abbreviated World Health Organization Quality of Life.

**Figure 3.** Estimated alcohol use and psychological distress scores over time, adjusted for gender, age, and binary measures of alcohol use disorder, any drug use, and mental health disorder (major depression, generalized anxiety, and posttraumatic stress disorder). HDD: heavy drinking days; SIP-R: Short Inventory of Problems–Revised.



## Discussion

### Principal Findings

The low rates of and delays in accessing treatment for alcohol-related problems among OEF and OIF veterans can detrimentally affect their health, long-term functioning, and reintegration into communities. While specialty substance use disorder care is seen as the gold standard for alcohol treatment, mHealth apps such as Step Away have great public health reach [15] and are a promising alternative for veterans who are unable or reluctant to seek traditional clinic-based alcohol treatment. This study contributes to the emerging literature on alcohol treatment apps and to the investigations of Step Away, the next iteration of LBMI-A, in particular. Results from this pilot study indicated that OEF and OIF veterans were willing to engage with Step Away; found the app acceptable with respect to effectiveness, efficiency, and overall satisfaction; and decreased their heavy drinking days over 6 months of study participation.

While veterans recruited for this study were selected based on reports of drinking over recommended limits, they were not required to have an AUD, be seeking care, or have a goal of changing their drinking behavior. Despite these facts, over 80% (45/55) of the sample met criteria for AUD and over one-third had received prior substance use treatment at baseline. Given our sample and the mixed findings from other studies of alcohol-related apps, many showing modest reductions in use [22,23], the decreases seen in heavy drinking days and alcohol-related problems scores are encouraging. Future studies

with a larger sample size and a control group are needed to validate these findings and determine their association with Step Away.

Weekly episodes of Step Away use among those who accessed the app at least once during a given week remained steady over the study course, but time per episode dropped quickly after week 1. Such findings are similar or better than those of other studies (eg, 5% of participants using the app at week 12 [22]). This drop-off in use is understandable; we anticipated that most participants would complete the most time-consuming aspects of Step Away (eg, setting up their preferences and goals and reading through psychoeducational material) during week 1. Sustained use consisted mainly of the daily and weekly feedback features of Step Away, which continued to be used by approximately one-third of participants through week 24. This may be driven by Step Away's use of push notifications, as recommended in the literature [15,20,50]. While these features are important, other components based in MET (eg, goal setting) and community reinforcement (eg, social supports) were not as well accessed.

This raises the question of how to deliver treatment components that are thought to be essential to quality substance use disorder care in the virtual environment. Some apps (eg, A-CHESS) are designed to be adjunctive to standard in-person care, either to be used as an additional component of care or a continuation of care. For stand-alone treatment apps such as Step Away, the issue of how to engage app users with evidence-based interventions is critical. Future iterations of Step Away must

leverage technology to deliver these interventions in a way that is perceived as relevant and useful. Personalizing app content (ie, tailoring) is associated with increases in user-rated quality and popularity [15]. Apps such as InDEx [23], which is geared specifically to a veteran population, included highly tailored messaging and saw high rates of app engagement, with two-thirds of participants using the app every week for 4 weeks after enrollment compared to 43.6% (24/55) seen in this study. More sophisticated use of tailoring that introduces new interventions specific to the app user's treatment course (eg, relapse prevention) might help to keep users fully engaged in app content and reengage those individuals whose app usage has dropped off. Combining Step Away with provider contact, peer support, or both, a recommended means of increasing engagement [51], is under study by Blonigen and colleagues [52] and will be critical to our understanding of how to best improve engagement with alcohol treatment apps.

Similar to findings pertaining to engagement with the DrinkLess app [24], which was tested in the general population, engagement with Step Away as measured by weeks of use and total minutes of use was positively associated with age, female gender, education level, and an income above US \$50,000. Examinations of the relationship between drinking characteristics and app engagement have yielded inconsistent results, with some studies identifying an association between high drinking levels and app engagement [22] and others detecting no association [24]. Our findings of no significant associations between baseline drinking severity and engagement must be considered in relation to our sample. The study was presented as an app evaluation rather than a treatment study, and as noted above, participants were not required to have a desire to cut back on alcohol use. It is also possible that engagement reflected a desire to be compliant with the study rather than treatment engagement per se. Our findings also do not help clarify whether app engagement improves drinking outcomes. While we did not detect an association between engagement and outcomes, our sample size was small and likely underpowered. Additional research, ideally with a comparison condition and a larger sample of participants desiring to cut back on use, is needed to determine whether Step Away is more appealing to veterans with alcohol misuse and no AUD compared with those with AUD and to assess the relationship between app engagement and outcomes.

### Study Limitations

While the primary aim of this study was to evaluate the usability of Step Away, we were not able to tease out the main

components of usability—effectiveness, efficiency, and satisfaction—because the SUS is a composite measure. In addition, we were unable to determine with accuracy the amount of time spent in each Step Away module for the first time. Thus, learnability was assessed by time spent in the app on the first episode of use and first week of use, with time spent in the app after week 1 provided for context. The study relies on veteran self-report for measuring alcohol use and prior receipt of alcohol-related care without verification by laboratory testing and treatment utilization records. Without a comparison condition, we do not know if reductions in use were due to use of Step Away or other factors; other studies of alcohol treatment apps [17] have found that reductions in use are modest when compared to controls and that differences decline over time. Furthermore, we did not assess participants' motivations for using the app or whether they desired or saw a need for cutting back on alcohol consumption. While we assessed general distress using the Kessler-10, we did not examine symptoms related to co-occurring mental health conditions that are common among veterans (eg, depression, posttraumatic stress disorder). As iPhone users tend to be less racially diverse and report higher incomes than other smartphone users, the sample of participants recruited for this study is not likely to have been representative of OEF and OIF veterans with AUD. Further, as Tofighi et al [51] point out in their recent review, access barriers to technology (eg, smartphone ownership, connectivity) often mirror the very barriers to treatment that apps are attempting to overcome. Future studies must consider how best to disseminate treatment apps to those who need them and how to overcome technology and connectivity barriers. Our sample was small and composed of non-treatment-seeking OEF and OIF veterans who were aged 18 to 55 years and enrolled in VA care; our findings may not generalize to other OEF and OIF veterans, veterans of other service eras, or individuals in the community.

This pilot study represents one of the first studies to evaluate the effects of an interactive mHealth app for alcohol misuse in a community sample of veterans. Participants' use of the app and alcohol-related outcomes were followed for 6 months, longer than many studies of apps addressing alcohol use. Acceptability and usability results provided essential and timely information to guide further refinement of Step Away (eg, modifications of text-heavy screens) and implementation of mHealth apps among veterans with alcohol misuse. If results are verified in a larger controlled trial, Step Away has the potential to improve access to alcohol-related care for a subset of the veteran population that is known to be reluctant to seek such care.

### Acknowledgments

This material is based upon work supported by the VA Health Services Research and Development, Merit Review Pilot Study (PPO 15-410-2), and the VA Center of Excellence in Substance Addiction Treatment and Education, Seattle, WA. Supporting organizations had no further role in the study design; collection, analysis, and interpretation of data; writing of the report; or decision to submit the paper for publication.

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the US Department of Veterans Affairs or University of Washington.

## Conflicts of Interest

PLD has an ownership interest in the company that owns Step Away.

## References

1. Room R, Babor T, Rehm J. Alcohol and public health. *Lancet* 2005 Feb 05;365(9458):519-530. [doi: [10.1016/S0140-6736\(05\)17870-2](https://doi.org/10.1016/S0140-6736(05)17870-2)] [Medline: [15705462](https://pubmed.ncbi.nlm.nih.gov/15705462/)]
2. Rehm J, Gmel G, Sempos CT, Trevisan M. Alcohol-related morbidity and mortality. *Alcohol Res Health* 2003;27(1):39-51 [FREE Full text] [Medline: [15301399](https://pubmed.ncbi.nlm.nih.gov/15301399/)]
3. Burnett-Zeigler I, Ilgen M, Valenstein M, Zivin K, Gorman L, Blow A, et al. Prevalence and correlates of alcohol misuse among returning Afghanistan and Iraq veterans. *Addict Behav* 2011 Aug;36(8):801-806. [doi: [10.1016/j.addbeh.2010.12.032](https://doi.org/10.1016/j.addbeh.2010.12.032)] [Medline: [21482030](https://pubmed.ncbi.nlm.nih.gov/21482030/)]
4. Calhoun PS, Elter JR, Jones ER, Kudler H, Straits-Tröster K. Hazardous alcohol use and receipt of risk-reduction counseling among U.S. veterans of the wars in Iraq and Afghanistan. *J Clin Psychiatry* 2008 Nov;69(11):1686-1693. [doi: [10.4088/jcp.v69n1103](https://doi.org/10.4088/jcp.v69n1103)] [Medline: [19012816](https://pubmed.ncbi.nlm.nih.gov/19012816/)]
5. Hawkins EJ, Lapham GT, Kivlahan DR, Bradley KA. Recognition and management of alcohol misuse in OEF/OIF and other veterans in the VA: A cross-sectional study. *Drug Alcohol Depend* 2010 Jun 01;109(1-3):147-153. [doi: [10.1016/j.drugalcdep.2009.12.025](https://doi.org/10.1016/j.drugalcdep.2009.12.025)] [Medline: [20167440](https://pubmed.ncbi.nlm.nih.gov/20167440/)]
6. Bradley KA, Ludman EJ, Chavez LJ, Bobb JF, Ruedebusch SJ, Achtmeyer CE, et al. Patient-centered primary care for adults at high risk for AUDs: The Choosing Healthier Drinking Options In primary Care (CHOICE) trial. *Addict Sci Clin Pract* 2017 May 17;12(1):15. [doi: [10.1186/s13722-017-0080-2](https://doi.org/10.1186/s13722-017-0080-2)] [Medline: [28514963](https://pubmed.ncbi.nlm.nih.gov/28514963/)]
7. Goldberg SB, Simpson TL, Lehavot K, Katon JG, Chen JA, Glass JE, et al. Mental health treatment delay: A comparison among civilians and veterans of different service eras. *Psychiatr Serv* 2019 May 01;70(5):358-366 [FREE Full text] [doi: [10.1176/appi.ps.201800444](https://doi.org/10.1176/appi.ps.201800444)] [Medline: [30841842](https://pubmed.ncbi.nlm.nih.gov/30841842/)]
8. Maguen S, Madden E, Cohen BE, Bertenthal D, Seal KH. Time to treatment among veterans of conflicts in Iraq and Afghanistan with psychiatric diagnoses. *Psychiatr Serv* 2012 Dec;63(12):1206-1212. [doi: [10.1176/appi.ps.201200051](https://doi.org/10.1176/appi.ps.201200051)] [Medline: [23070131](https://pubmed.ncbi.nlm.nih.gov/23070131/)]
9. Willenbring ML. Gaps in clinical prevention and treatment for alcohol use disorders: Costs, consequences, and strategies. *Alcohol Res* 2013;35(2):238-243 [FREE Full text] [Medline: [24881332](https://pubmed.ncbi.nlm.nih.gov/24881332/)]
10. Hoge CW, Castro CA, Messer SC, McGurk D, Cotting DI, Koffman RL. Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *N Engl J Med* 2004 Jul 01;351(1):13-22. [doi: [10.1056/NEJMoa040603](https://doi.org/10.1056/NEJMoa040603)] [Medline: [15229303](https://pubmed.ncbi.nlm.nih.gov/15229303/)]
11. Stecker T, Shiner B, Watts BV, Jones M, Conner KR. Treatment-seeking barriers for veterans of the Iraq and Afghanistan conflicts who screen positive for PTSD. *Psychiatr Serv* 2013 Mar;64(3):280-283. [doi: [10.1176/appi.ps.001372012](https://doi.org/10.1176/appi.ps.001372012)] [Medline: [23450385](https://pubmed.ncbi.nlm.nih.gov/23450385/)]
12. Dulin PL, Gonzalez VM, King DK, Giroux D, Bacon S. Smartphone-based, self-administered intervention system for alcohol use disorders: Theory and empirical evidence basis. *Alcohol Treat Q* 2013;31(3):321-336 [FREE Full text] [doi: [10.1080/07347324.2013.800425](https://doi.org/10.1080/07347324.2013.800425)] [Medline: [24347811](https://pubmed.ncbi.nlm.nih.gov/24347811/)]
13. 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](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
14. Mobile fact sheet. Pew Research Center. 2019. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2020-03-19]
15. Hoepfner BB, Schick MR, Kelly LM, Hoepfner SS, Bergman B, Kelly JF. There is an app for that - Or is there? A content analysis of publicly available smartphone apps for managing alcohol use. *J Subst Abuse Treat* 2017 Nov;82:67-73. [doi: [10.1016/j.jsat.2017.09.006](https://doi.org/10.1016/j.jsat.2017.09.006)] [Medline: [29021117](https://pubmed.ncbi.nlm.nih.gov/29021117/)]
16. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: Content analysis. *J Med Internet Res* 2015;17(5):e118 [FREE Full text] [doi: [10.2196/jmir.4060](https://doi.org/10.2196/jmir.4060)] [Medline: [25977135](https://pubmed.ncbi.nlm.nih.gov/25977135/)]
17. Gajecki M, Andersson C, Rosendahl I, Sinadinovic K, Fredriksson M, Berman AH. Skills training via smartphone app for university students with excessive alcohol consumption: A randomized controlled trial. *Int J Behav Med* 2017 Dec;24(5):778-788 [FREE Full text] [doi: [10.1007/s12529-016-9629-9](https://doi.org/10.1007/s12529-016-9629-9)] [Medline: [28224445](https://pubmed.ncbi.nlm.nih.gov/28224445/)]
18. Hides L, Quinn C, Cockshaw W, Stoyanov S, Zelenko O, Johnson D, et al. Efficacy and outcomes of a mobile app targeting alcohol use in young people. *Addict Behav* 2018 Feb;77:89-95. [doi: [10.1016/j.addbeh.2017.09.020](https://doi.org/10.1016/j.addbeh.2017.09.020)] [Medline: [28992580](https://pubmed.ncbi.nlm.nih.gov/28992580/)]
19. Cohn AM, Hunter-Reel D, Hagman BT, Mitchell J. Promoting behavior change from alcohol use through mobile technology: The future of ecological momentary assessment. *Alcohol Clin Exp Res* 2011 Dec;35(12):2209-2215 [FREE Full text] [doi: [10.1111/j.1530-0277.2011.01571.x](https://doi.org/10.1111/j.1530-0277.2011.01571.x)] [Medline: [21689119](https://pubmed.ncbi.nlm.nih.gov/21689119/)]
20. Dedert E, Williams JW, Stein R, McNeil JM, McDuffie J, Ross I. e-Interventions for alcohol misuse. US Department of Veterans Affairs. 2014. URL: <https://www.ncbi.nlm.nih.gov/books/NBK293699/> [accessed 2021-03-19]

21. Connolly SL, Miller CJ, Koenig CJ, Zamora KA, Wright PB, Stanley RL, et al. Veterans' attitudes toward smartphone app use for mental health care: Qualitative study of rurality and age differences. *JMIR Mhealth Uhealth* 2018 Aug 22;6(8):e10748 [FREE Full text] [doi: [10.2196/10748](https://doi.org/10.2196/10748)] [Medline: [30135050](https://pubmed.ncbi.nlm.nih.gov/30135050/)]
22. Attwood S, Parke H, Larsen J, Morton KL. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. *BMC Public Health* 2017 May 17;17(1):394 [FREE Full text] [doi: [10.1186/s12889-017-4358-9](https://doi.org/10.1186/s12889-017-4358-9)] [Medline: [28511698](https://pubmed.ncbi.nlm.nih.gov/28511698/)]
23. Leightley D, Puddephatt J, Jones N, Mahmoodi T, Chui Z, Field M, et al. A smartphone app and personalized text messaging framework (InDEx) to monitor and reduce alcohol use in ex-serving personnel: Development and feasibility study. *JMIR Mhealth Uhealth* 2018 Sep 11;6(9):e10074 [FREE Full text] [doi: [10.2196/10074](https://doi.org/10.2196/10074)] [Medline: [30206054](https://pubmed.ncbi.nlm.nih.gov/30206054/)]
24. Garnett C, Perski O, Tombor I, West R, Michie S, Brown J. Predictors of engagement, response to follow up, and extent of alcohol reduction in users of a smartphone app (Drink Less): Secondary analysis of a factorial randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Dec 14;6(12):e11175 [FREE Full text] [doi: [10.2196/11175](https://doi.org/10.2196/11175)] [Medline: [30552081](https://pubmed.ncbi.nlm.nih.gov/30552081/)]
25. Dulin PL, Gonzalez VM, Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: Usability and early outcomes. *Subst Abus* 2014;35(2):168-175 [FREE Full text] [doi: [10.1080/08897077.2013.821437](https://doi.org/10.1080/08897077.2013.821437)] [Medline: [24821354](https://pubmed.ncbi.nlm.nih.gov/24821354/)]
26. Gonzalez VM, Dulin PL. Comparison of a smartphone app for alcohol use disorders with an internet-based intervention plus bibliotherapy: A pilot study. *J Consult Clin Psychol* 2015 Apr;83(2):335-345. [doi: [10.1037/a0038620](https://doi.org/10.1037/a0038620)] [Medline: [25622202](https://pubmed.ncbi.nlm.nih.gov/25622202/)]
27. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: A randomized clinical trial. *JAMA Psychiatry* 2014 May;71(5):566-572 [FREE Full text] [doi: [10.1001/jamapsychiatry.2013.4642](https://doi.org/10.1001/jamapsychiatry.2013.4642)] [Medline: [24671165](https://pubmed.ncbi.nlm.nih.gov/24671165/)]
28. Ergonomic requirements for office work with visual display terminals (VDTs) Part 11 Guidance on usability. International standard, 9241-11. International Organization for Standardization. 1998. URL: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:en> [accessed 2021-03-19]
29. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: A systematic literature review. *J Med Syst* 2015 Feb;39(2):1. [doi: [10.1007/s10916-014-0182-2](https://doi.org/10.1007/s10916-014-0182-2)] [Medline: [25600193](https://pubmed.ncbi.nlm.nih.gov/25600193/)]
30. Bush K, Kivlahan DR, McDonnell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking. *Arch Intern Med* 1998 Sep 14;158(16):1789-1795 [FREE Full text] [doi: [10.1001/archinte.158.16.1789](https://doi.org/10.1001/archinte.158.16.1789)] [Medline: [9738608](https://pubmed.ncbi.nlm.nih.gov/9738608/)]
31. Blonigen DM, Harris-Olenak B, Haber JR, Kuhn E, Timko C, Humphreys K, et al. Customizing a clinical app to reduce hazardous drinking among veterans in primary care. *Psychol Serv* 2019 May;16(2):250-254. [doi: [10.1037/ser0000300](https://doi.org/10.1037/ser0000300)] [Medline: [30407060](https://pubmed.ncbi.nlm.nih.gov/30407060/)]
32. Enhancing motivation for change in substance abuse treatment. Treatment Improvement Protocol (TIP) Series, No. 35. Center for Substance Abuse Treatment. Rockville, MD: Substance Abuse and Mental Health Services Administration; 1999. URL: <https://www.ncbi.nlm.nih.gov/books/NBK64967/> [accessed 2021-03-19]
33. Marlatt G, Donovan D. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. 2nd ed. New York, NY: Guilford Press; 2005.
34. Meyers RJ, Villanueva M, Smith JE. The community reinforcement approach: History and new directions. *J Cogn Psychother* 2005 Jul 01;19(3):247-260. [doi: [10.1891/jcop.2005.19.3.247](https://doi.org/10.1891/jcop.2005.19.3.247)]
35. Smith J, Meyers R. Motivating Substance Abusers to Enter Treatment: Working With Families. New York, NY: Guilford Press; 2004.
36. D'Zurilla T, Nezu A. Problem-Solving Therapy: A Positive Approach to Clinical Intervention. 3rd ed. New York, NY: Spring Publishing; 2007.
37. 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):22-33. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
38. Brooke J. SUS: A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London, UK: Taylor and Francis; 1996:189-194.
39. Bangor A, Kortum PT, Miller JT. Determining what individual SUS scores mean: Adding an adjective rating scale. *J Usability Stud* 2009;4(3):114-123 [FREE Full text]
40. Lewis JR, Sauro J. The factor structure of the System Usability Scale. In: Kurosu M, editor. Human Centered Design: Lecture Notes in Computer Science, vol 5619. Berlin, Germany: Springer; 2009:94-103.
41. Lindgaard G, Fernandes G, Dudek C, Brown J. Attention web designers: You have 50 milliseconds to make a good first impression!. *Behav Inf Technol* 2006 Mar;25(2):115-126. [doi: [10.1080/01449290500330448](https://doi.org/10.1080/01449290500330448)]
42. Bevan N. Measuring usability as quality of use. *Software Qual J* 1995 Jun;4(2):115-130. [doi: [10.1007/BF00402715](https://doi.org/10.1007/BF00402715)]
43. Sobell LC, Sobell MB. Timeline Followback: A technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen J, editors. Measuring Alcohol Consumption: Psychosocial and Biological Methods. Totowa, NJ: Humana Press; 1992:41-72.

44. Kiluk BD, Dreifuss JA, Weiss RD, Morgenstern J, Carroll KM. The Short Inventory of Problems – Revised (SIP-R): Psychometric properties within a large, diverse sample of substance use disorder treatment seekers. *Psychol Addict Behav* 2013;27(1):307-314. [doi: [10.1037/a0028445](https://doi.org/10.1037/a0028445)] [Medline: [22642856](https://pubmed.ncbi.nlm.nih.gov/22642856/)]
45. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). *Aust N Z J Public Health* 2001 Dec;25(6):494-497. [doi: [10.1111/j.1467-842x.2001.tb00310.x](https://doi.org/10.1111/j.1467-842x.2001.tb00310.x)] [Medline: [11824981](https://pubmed.ncbi.nlm.nih.gov/11824981/)]
46. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002 Aug;32(6):959-976. [doi: [10.1017/s0033291702006074](https://doi.org/10.1017/s0033291702006074)] [Medline: [12214795](https://pubmed.ncbi.nlm.nih.gov/12214795/)]
47. Zubaran C, Foresti K. Quality of life and substance use: Concepts and recent tendencies. *Curr Opin Psychiatry* 2009;22(3):281-286. [doi: [10.1097/ycp.0b013e328328d154](https://doi.org/10.1097/ycp.0b013e328328d154)] [Medline: [19378380](https://pubmed.ncbi.nlm.nih.gov/19378380/)]
48. International guide for monitoring alcohol consumption and related harm. World Health Organization. Copenhagen, Denmark: World Health Organization, Department of Mental Health and Substance Dependence, Noncommunicable Diseases and Mental Health Cluster; 2000. URL: <https://apps.who.int/iris/handle/10665/66529> [accessed 2021-03-19]
49. Greenfield TK. Evaluating competing models of alcohol-related harm. *Alcohol Clin Exp Res* 1998 Apr;22(Suppl 2):52S-62S. [doi: [10.1097/00000374-199802001-00008](https://doi.org/10.1097/00000374-199802001-00008)] [Medline: [9603307](https://pubmed.ncbi.nlm.nih.gov/9603307/)]
50. Song T, Qian S, Yu P. Mobile health interventions for self-control of unhealthy alcohol use: Systematic review. *JMIR Mhealth Uhealth* 2019 Jan 29;7(1):e10899 [FREE Full text] [doi: [10.2196/10899](https://doi.org/10.2196/10899)] [Medline: [30694200](https://pubmed.ncbi.nlm.nih.gov/30694200/)]
51. Tofighi B, Abrantes A, Stein MD. The role of technology-based interventions for substance use disorders in primary care. *Med Clin North Am* 2018 Jul;102(4):715-731. [doi: [10.1016/j.mcna.2018.02.011](https://doi.org/10.1016/j.mcna.2018.02.011)] [Medline: [29933825](https://pubmed.ncbi.nlm.nih.gov/29933825/)]
52. Blonigen D, Harris-Olenak B, Kuhn E, Humphreys K, Timko C, Dulin P. From "Step Away" to "Stand Down": Tailoring a smartphone app for self-management of hazardous drinking for veterans. *JMIR Mhealth Uhealth* 2020 Feb 13;8(2):e16062 [FREE Full text] [doi: [10.2196/16062](https://doi.org/10.2196/16062)] [Medline: [32053118](https://pubmed.ncbi.nlm.nih.gov/32053118/)]

## Abbreviations

**A-CHESS:** Addiction Comprehensive Health Enhancement Support System  
**AUD:** alcohol use disorder  
**DSM-5:** Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision  
**HDD:** heavy drinking days  
**LBMI-A:** Location-Based Monitoring and Intervention System for Alcohol Use Disorders  
**MET:** motivational enhancement therapy  
**mHealth:** mobile health  
**MINI:** Mini International Neuropsychiatric Interview  
**OEF:** Operation Enduring Freedom  
**OIF:** Operation Iraqi Freedom  
**SIP-R:** Short Inventory of Problems–Revised  
**SUS:** System Usability Scale  
**TLFB:** Timeline Followback  
**VA:** Veterans Affairs  
**WHOQOL-BREF:** abbreviated World Health Organization Quality of Life

*Edited by L Buis; submitted 20.11.20; peer-reviewed by D Leightley, K Yokotani; comments to author 08.12.20; revised version received 02.02.21; accepted 10.03.21; published 08.04.21.*

*Please cite as:*

Malte CA, Dulin PL, Baer JS, Fortney JC, Danner AN, Lott AMK, Hawkins EJ

*Usability and Acceptability of a Mobile App for the Self-Management of Alcohol Misuse Among Veterans (Step Away): Pilot Cohort Study*

*JMIR Mhealth Uhealth* 2021;9(4):e25927

URL: <https://mhealth.jmir.org/2021/4/e25927>

doi: [10.2196/25927](https://doi.org/10.2196/25927)

PMID: [33830064](https://pubmed.ncbi.nlm.nih.gov/33830064/)

©Carol A Malte, Patrick L Dulin, John S Baer, John C Fortney, Anissa N Danner, Aline M K Lott, Eric J Hawkins. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 08.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Mobile Phone Intervention Based on an HIV Risk Prediction Tool for HIV Prevention Among Men Who Have Sex With Men in China: Randomized Controlled Trial

Ke Yun<sup>1,2,3,4</sup>, PhD; Zhenxing Chu<sup>1,2,3,4</sup>, MD; Jing Zhang<sup>1,2,3,4</sup>, PhD; Wenqing Geng<sup>1,2,3,4</sup>, PhD; Yongjun Jiang<sup>1,2,3,4</sup>, PhD; Willa Dong<sup>5</sup>, MSPH; Hong Shang<sup>1,2,3,4\*</sup>, PhD; Junjie Xu<sup>1,2,3,4\*</sup>, PhD

<sup>1</sup>NHC Key Laboratory of AIDS Immunology (China Medical University), National Clinical Research Center for Laboratory Medicine, The First Affiliated Hospital of China Medical University, Shenyang, China

<sup>2</sup>Key Laboratory of AIDS Immunology, Chinese Academy of Medical Sciences, Shenyang, China

<sup>3</sup>Key Laboratory of AIDS Immunology of Liaoning Province, Shenyang, China

<sup>4</sup>Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China

<sup>5</sup>Department of Health Behavior, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

\*these authors contributed equally

**Corresponding Author:**

Hong Shang, PhD

NHC Key Laboratory of AIDS Immunology (China Medical University)

National Clinical Research Center for Laboratory Medicine

The First Affiliated Hospital of China Medical University

No 155, Nanjing North Street, Heping District

Shenyang

China

Phone: 86 8328 2634

Email: [hongshang100@hotmail.com](mailto:hongshang100@hotmail.com)

## Abstract

**Background:** eHealth interventions based on risk stratification have not been extensively applied for HIV behavioral interventions among HIV-negative men who have sex with men (MSM).

**Objective:** This study aimed to evaluate the efficacy of a mobile phone intervention based on an HIV risk prediction tool in promoting HIV testing and reducing high-risk behavior among HIV-negative MSM in China.

**Methods:** We performed a mobile phone-based randomized controlled clinical trial for 12 weeks. A comprehensive intervention package deployed on Jinshuju—an online survey platform—was developed and consisted of 4 components: (1) a validated HIV risk prediction tool that provides information on personalized risk reduction interventions; (2) a map of individualized HIV testing facilities based on their geographic location; (3) a QR code for free resources on HIV prevention, including condoms and HIV self-testing kits; and (4) general resources for HIV health education. MSM participants recruited from WeChat/QQ groups were randomly assigned to the intervention or control group at a 1:1 ratio. The staff sent the QR code for the comprehensive intervention package to MSM in the intervention group over WeChat and sent the QR code only for the resources on HIV health education to those in the control group. At baseline and 12-week follow-up, data on HIV-related risk behavior and HIV testing behavior were collected through the Jinshuju online survey platform.

**Results:** In total, 192 MSM were recruited and assigned to the intervention or control group (n=96 each). At week 12, the total clinical trial retention rate was 87.5%. The number of male sexual partners of the MSM in the past 3 months was significantly lower in the intervention group than in the control group (3.51, SD 4.1 vs 6.01, SD 11.4, respectively; mean difference -2.5; 95% CI -5.12 to 0.12;  $P=.05$ ); the rate of condom use with casual sexual partners was higher in the intervention group than in the control group (87%, n=66/76 vs 70%, n=54/77 respectively; odds ratio 2.81, 95% CI 1.23-6.39;  $P=.01$ ). The proportion of individuals intending to undergo HIV testing after in the following 30 days was marginally higher in the intervention group than in the control group (90%, n=77/86 vs 79%, n=65/82 respectively; odds ratio 2.20, 95% CI 0.90-5.35;  $P=.07$ ). The incremental cost-effectiveness ratio of eHealth intervention was US \$131.60 on reducing 1 sexual partner and US \$19.70 for a 1% increment in condom usage with casual partners.

**Conclusions:** A comprehensive intervention based on an HIV risk prediction tool can reduce the number of male sexual partners among MSM and increase the rate of condom use with casual partners. Hence, this intervention is a very promising preventive strategy for HIV among MSM, especially in areas with a prominent HIV epidemic.

**Trial Registration:** Chinese Clinical Trial Registry ChiCTR1800017268; <http://www.chictr.org.cn/showprojen.aspx?proj=29271>

(*JMIR Mhealth Uhealth* 2021;9(4):e19511) doi:[10.2196/19511](https://doi.org/10.2196/19511)

## KEYWORDS

eHealth intervention; high-risk behavior intervention; HIV risk prediction; men who have sex with men

## Introduction

Men who have sex with men (MSM) are a significant target population for the prevention and control of HIV infections worldwide [1]. In China, although traditional sexual behavioral interventions, such as promoting abstinence and condom use, have played an essential role in curbing the HIV epidemic, HIV acquisition among MSM, especially young MSM, remains poorly controlled [2]. In 2015, HIV prevalence among MSM was 8%, and its incidence was 5.61 cases per 100 person-years [3]. Young MSM aged 20-29 years may be disproportionately affected because they accounted for 47.0% of the 120,371 HIV cases recorded during 2006-2015 in China [4].

In recent years, internet-based dating websites or smartphone apps have been increasingly used by MSM to seek sexual partners [5]. In the United States, a study reported that among 195 MSM aged 19-24 years, 75% of participants reported having sexual encounters with individuals through these apps [6]. Similarly, in China, out of 353 MSM, almost 70% reported having sexual experiences with men they met on the internet [7]. Furthermore, among 375 young MSM, 70% reported having unprotected anal intercourse with partners from the internet-based dating app Grindr and perceived a low risk of HIV acquisition [8]. Fortunately, internet-based eHealth interventions have emerged as prominent tools for preventing HIV acquisition among MSM owing to their convenience and low cost [9]. Therefore, considering the popularity of internet-based platforms for interacting with sexual partners, social media may be leveraged to deliver behavioral interventions, especially among young MSM, who are not only particularly vulnerable to an HIV infection compared to those in other age groups [10] but also willing to receive HIV prevention interventions via smartphone apps [11].

We previously developed and validated an HIV risk prediction model and constructed a social media platform-based HIV risk

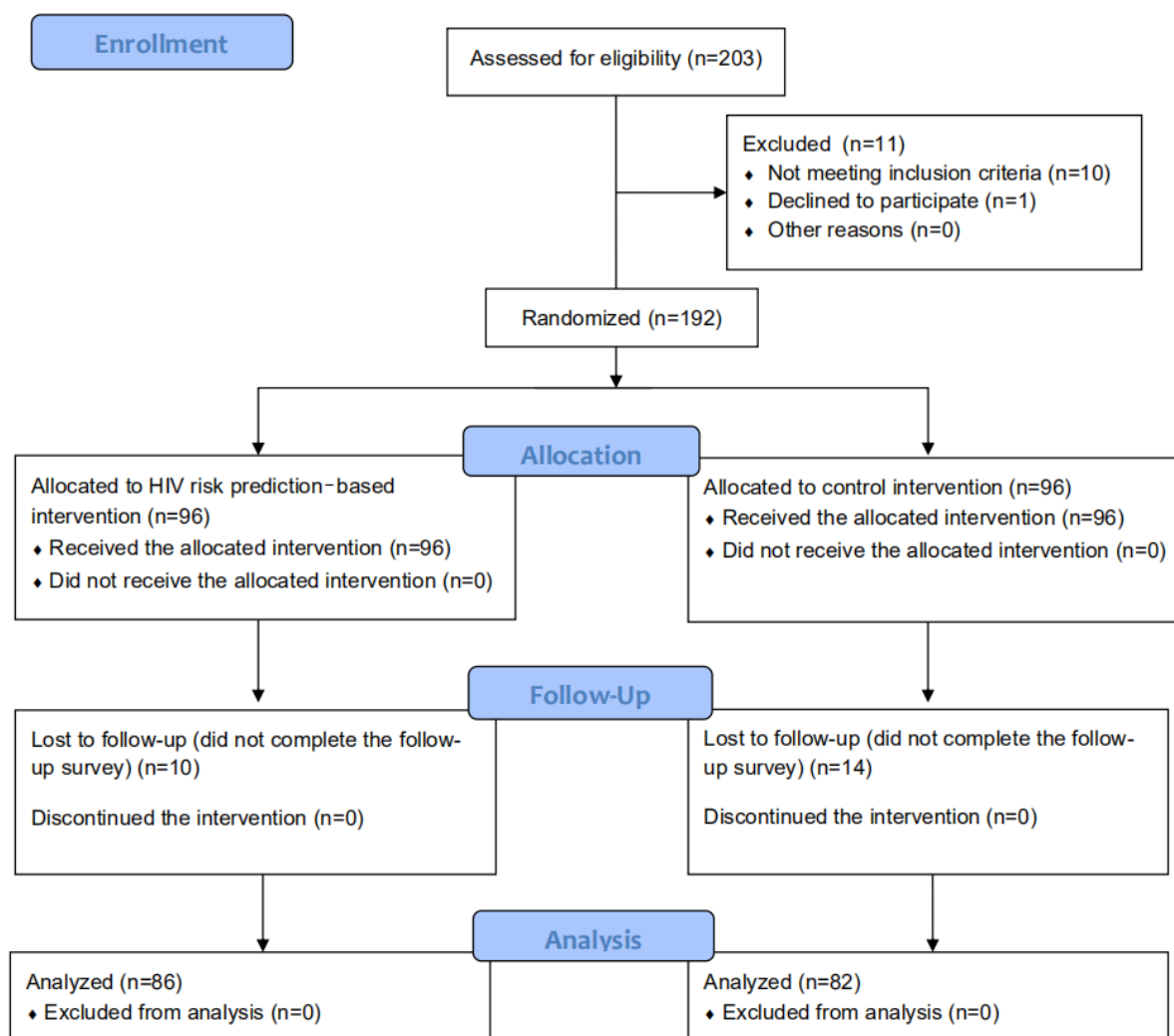
assessment tool [12]. Based on the HIV risk assessment tool, we then developed a comprehensive intervention package for MSM to estimate their probability of acquiring an HIV infection and to provide personalized and tailored information on HIV testing and reducing behavioral risks. This study aimed to assess the preliminary efficacy of an eHealth intervention, based on a previously developed and validated HIV risk prediction tool [12], in reducing the HIV risk and promoting HIV testing among MSM in China.

## Methods

### Study Population and Design

This 2-arm, parallel, randomized, double-blinded clinical trial was conducted in accordance with the Consolidated Standard of Reporting Trials guidelines [13] for randomized controlled trials ([Multimedia Appendix 1](#)) and the Checklist for Reporting Results of Internet E-Surveys [14] for online surveys ([Multimedia Appendix 2](#)). [Figure 1](#) shows a Consolidated Standard of Reporting Trials flow diagram of the study design. From October 2017 to March 2018, participants were recruited through an advertisement on a popular WeChat/QQ group for MSM in China with the assistance of Shenyang Sunshine Working Group, a local MSM community-based organization (CBO). Following the completion of a web-based screening questionnaire, the computer algorithm immediately provided an eligibility assessment. The inclusion criteria were as follows: men who had anal or oral intercourse with men in the previous year, had not been diagnosed with an HIV infection, owned an Android or Apple smartphone and a WeChat account, were aged  $\geq 18$  years, provided written informed consent, and were able to comprehend written Chinese (Mandarin). The exclusion criteria were as follows: not having a WeChat account and having tested positive for HIV. Informed consent was received from all MSM involved in the survey.

**Figure 1.** Consolidated Standards of Reporting Trials flowchart for the recruitment of participants who are men who have sex with men in Shenyang, China.



## Randomization and Study Procedures

All eligible MSM were randomly assigned to an intervention group and a control group on Microsoft Excel 2010 at a 1:1 ratio after obtaining informed consent. Participants and investigators were both blinded to group allocation. At baseline and prospective follow-up, MSM included in our study received a push notification with a QR code for a link to the questionnaire through the study's WeChat account. These 28-item questionnaires were displayed on a webpage pertaining to high-risk sexual behavior and HIV testing in the previous 3 months and contained questions on the number of sexual partners, condom usage, practices related to unprotected anal intercourse, and intentions to undergo HIV testing in the following 30 days. Before administering the questionnaire in an open-field survey, a small-scale pilot study was conducted to ensure the reliability and validity of the survey instrument. Participants were required to answer all questions, and we programmed skip patterns, which are a feature of the survey platform, to reduce participant burden. The respondents could review and change their answers before submission. Participants who completed the web-based questionnaire survey received a ¥20 (approximately US \$3) subsidy to thank them for their

participation. Implementation was carried out as follows. First, after the completion of the recruitment screening questionnaire, participants were asked to add the study's WeChat account as an active friend and then send screenshots of the completed questionnaire to the study WeChat account. After the staff assessed the integrity of the questionnaire, they transferred the subsidy to the participants through WeChat transfer. The staff used the WeChat nickname and WeChat ID to ensure that the questionnaire responders and recipients are the same individuals. Every time the participants completed the questionnaire, they received the subsidy through the app.

## Intervention

The theoretical framework for the comprehensive intervention package used in our study is the AIDS Risk Reduction Model established in 1990 [15]. The staff sent the link to the QR code for the comprehensive intervention package to MSM in the intervention group over WeChat, and a link to a QR code for a webpage with resources on HIV/AIDS health education was sent to MSM in the control group. The intervention package included 4 components that were organized by the intervention index interface deployed on the Jinshuju data management platform, which included the following: (1) an HIV risk

prediction tool used for personalized evaluation of the HIV risk [12], feedback on risk behavior, and tailored suggestions for risk reduction; (2) recommendations for HIV testing facilities based on the user's location, which were churned by the questionnaire system backend after participants authorized the app to access their geographic location data; (3) web-based forms to receive order-free prevention resources on demand via mail, including condoms and HIV self-testing kits; and (4) an HIV health education website that contains information on HIV and associated hazards, the situation of the epidemic among MSM in China, and self-protection against HIV. All the information related to the intervention was presented on an external website. Detailed screenshots of the intervention interface are presented in [Multimedia Appendices 3-7](#). Every 4 weeks, information related to the intervention was sent again as a reminder to participants in both groups. Throughout the study period, MSM used WeChat as they did before the study.

To ensure the accuracy of the questionnaire survey, 3 measures were taken to avoid duplicate recruitment of the same individual in the study. First, participants could only complete the deployed questionnaire survey through WeChat, and secondary sharing of questionnaire links was prohibited. The questionnaire and data management system we used in this study restricted access to the specific survey pages in accordance with the IP address. Second, although WeChat nicknames are variable and dynamic, each WeChat account has a unique ID. On visiting the questionnaire, the survey system asked participants to authorize access to their personal WeChat ID information, and the unique WeChat ID was stored together with the survey results. Consequently, we could identify and verify the uniqueness of individuals by their WeChat ID. We did not use cookies to assign a unique user identifier to each client computer. Third, in the data processing stage, we deleted duplicate records from the same ID and records with identical responses, such as personal information and reported high-risk behavior. The first entry was retained for analysis.

### Primary Outcomes and Related Definitions

The primary outcomes of this study were the proportion of HIV testing, condom usage, anal intercourse in the past 3 months, and intention to undergo HIV testing in the next 30 days. The secondary outcome was the number of male sexual partners in the past 3 months. These outcomes were measured with a self-reported web-based questionnaire at 12 weeks post baseline. We defined an interval 7 days before and after the follow-up time point as the effective survey window for the follow-up time point. The observation period was determined in accordance with the survey experience of sexual behavioral among the MSM included in our study and previous randomized controlled trials [9]. Study subjects who failed to participate in the follow-up questionnaire during the survey window were considered as being lost to follow-up. The rate of being lost to follow-up was defined as the ratio of subjects who failed to participate in the 12-week follow-up to the total number of MSM recruited at baseline. Adherence was assessed by the percentage of individuals who visited the intervention index webpage (based on their IP addresses) out of the total number of participants in the intervention and control groups who completed the baseline questionnaire.

### Sample Size Calculation

We assumed that the intervention would be more effective than the resources on HIV/AIDS health education provided to the control group. According to our pilot survey, the assumption parameters were as follows: the average number of sexual partners in the past 3 months (primary study outcome) was 7 in the control group and 5 in the intervention group, and the SD for sample size estimation was 5. Other specific parameters are listed as follows: the degree of certainty ( $1-\beta$ ) was 80%, the significance level (Cronbach  $\alpha$ ) was .05, and 20% of subjects were lost to follow-up. Therefore, we intended to recruit 200 subjects. The sample size was estimated using the independent samples *t* test module of PASS 2008 (NCSS LLC).

### Cost Measurement and Cost-Effectiveness Criteria

The cost was estimated from the social perspective, and only the direct cost associated with the intervention was considered. The cost, including human capital cost, CBO recruitment and referral fees, communication fees, electricity fees, cost of condoms and lubricants, cost of HIV self-testing strips, postage fees, cost of the online questionnaire system, compensation for the participants, and stationary, was estimated in accordance with the clinical trial. The cost-effectiveness ratio (CER) and incremental cost-effectiveness ratio (ICER) were calculated as the economic evaluation indicators. The World Health Organization cost-effectiveness criteria were used for economic analysis, which state that interventions with an ICER of  $<0$  are effective and cost-effective, those with an ICER less than the average per capita GDP for a particular country or region are considered highly cost-effective, those with an ICER of  $<3$ -fold average per capita GDP are still considered cost-effective, and those that exceed this level are considered non-cost-effective [16]. The average per capita GDP for China in 2017 was US \$8497, which was considered the threshold indicating the cost-effectiveness of the intervention in this study.

### Statistical Analysis

Data were collected using the Jinshuju online survey platform (<https://jinshuju.net>).

Only completed eligible questionnaires were analyzed. Participants were required to answer all questions to minimize missing data and the need for imputation. Participants lost to follow-up were not included in the final intervention effect analysis. The chi-square test or Fisher exact probability test was used to compare categorical variables, and the independent samples *t* test was used to compare continuous data. All statistical analyses were performed using SAS (version 9.4, SAS Institute). Findings were considered significant when  $P < .05$ .

### Ethics Approval and Informed Consent

The Ethics Review Committee of the First Affiliated Hospital of China Medical University (Shenyang, China) approved the study protocol, investigation procedures, and questionnaires (approval# 2018-175-2). Written informed consent was obtained from all subjects before their participation in the questionnaire survey. WeChat IDs and self-reported dates of birth were used for individual identification to ensure that the participants' privacy was effectively protected.

## Results

### Recruitment, Screening, and Prospective Follow-up of Subjects

Between October 2017 and March 2018, of the 587 MSM who clicked on the recruitment advertisement, 203 completed the final screening questionnaire, and 192 were eligible for randomization (Figure 1). The mean age of the study subjects included was 28.5 (SD 7.9) years in the intervention group and 26.9 (SD 8.1) years in the control group. Among them, 113 (58.9%) had a Shenyang household registration, 114 (59.4%) had a monthly income of >¥3000 (approximately US \$437), 95

(49.5%) were single; 113 (58.9%) had a high school or higher degree, 74 (39.4%) were employed, 112 (58.3%) looked for sexual partners on the internet or on social media, and 173 (90.1%) had their first sexual experience at an age of <30 years. In the past 3 months, 176 (91.7%) participants exhibited same-sex sexual behavior, 120 (62.5%) had at least 2 male sexual partners, 103 (53.6%) had practiced unprotected anal intercourse, 77 (40.1%) had unprotected receptive anal intercourse, and 82 (42.7%) engaged in group sex. We observed no significant difference in the aforementioned baseline characteristics between the intervention group and control groups (Table 1).

**Table 1.** Baseline sociodemographic and behavioral characteristics of men who have sex with men in Shenyang, China, in the clinical trial (N=192).

Variable	Intervention group (n=96)	Control group (n=96)	P value
Age (years), mean (SD)	28.5 (7.9)	26.9 (8.1)	.18
<b>Local residents, n (%)</b>			.30
Yes	60 (63)	53 (55)	
No	36 (38)	43 (45)	
<b>Monthly income ( ¥ ), n (%)</b>			.14
≤3000	34 (35)	44 (46)	
>3000	62 (65)	52 (54)	
<b>Marital status, n (%)</b>			.67
Single	49 (51)	46 (48)	
Married or cohabiting with a partner	47 (49)	50 (52)	
<b>Education level, n (%)</b>			.95
Below high school	40 (42)	37 (39)	
High school	51 (53)	52 (55)	
College and above	5 (5)	5 (5)	
<b>Occupation, n (%)</b>			.64
Worker/staff	38 (40)	36 (39)	
Business	35 (37)	27 (29)	
Student	15 (16)	19 (21)	
Other	8 (8)	10 (11)	
<b>Age at first sexual experience (years), n (%)</b>			.22
<30	84 (88)	89 (93)	
≥30	12 (13)	7 (7)	
<b>Main venue to seek sexual partners, n (%)</b>			.42
Smartphone apps/web-based dating platforms	51 (53)	61 (64)	
Park/public bath/public toilet	6 (6)	4 (4)	
Bar/club	3 (3)	1 (1)	
Other	36 (38)	30 (31)	
<b>Displaying same-sex sexual behavior in the past 3 months, n (%)</b>			.30
Yes	90 (94)	86 (90)	
No	6 (6)	10 (10)	
<b>Having ≥2 male sexual partners in the past 3 months, n (%)</b>			.77
Yes	61 (64)	59 (62)	
No	35 (37)	37 (39)	
<b>Having unprotected anal intercourse in the past 3 months, n (%)</b>			.11
Yes	57 (59)	46 (48)	
No	39 (41)	50 (52)	
<b>Having unprotected receptive anal intercourse in the past 3 months, n (%)</b>			.66
Yes	40 (42)	37 (39)	
No	56 (58)	59 (62)	
<b>Engaging in group sex in the past 3 months, n (%)</b>			.70
Yes	45 (47)	37 (39)	
No	51 (53)	59 (62)	

Variable	Intervention group (n=96)	Control group (n=96)	P value
Number of male sexual partners in the past 3 months, mean (SD)	3.3 (3.4)	3.9 (4.7)	.33

### Adherence to the Intervention

In the past 3 months, the rate of page visits in the intervention group remained stable (68%, 52%, and 65% in months 1, 2, and 3, respectively), whereas those in the control group displayed an apparent downward trend (60%, 39%, and 18% in months 1, 2, and 3, respectively). For the 3 intervention modules in the intervention group, the risk assessment module had the highest page click rate.

### Efficacy of the Intervention

At week 12 of prospective follow-up, 168 questionnaires were collected (86 and 82 in the intervention and control groups, respectively), and the clinical trial retention rates were 90% (n=86/96) and 85% (n=82/96) for the intervention and control groups, respectively.

We observed no significant difference in the proportion of MSM who underwent HIV testing in the past 3 months. The proportion of MSM who intended to undergo HIV testing in the following 30 days was slightly higher in the intervention group than in the control group (90%, n=77/86 vs 79%, n=65/82, respectively; odds ratio [OR] 2.20, 95% CI 0.90-5.35;  $P=.07$ ). Condom usage among casual sexual partners in the past 3 months was significantly higher in the intervention group than in the control group (87%, n=66/76 vs 70%, n=54/77, respectively; OR 2.81, 95% CI 1.23-6.39;  $P=.01$ ). We observed no significant difference in the proportion of participants engaging in passive anal intercourse with or without a condom in the past 3 months and those engaging group sex. The number of male sexual partners in the past 3 months was significantly lower in the intervention group than in the control group (3.51, SD 4.1 vs 6.01, SD 11.4, respectively; mean difference=-2.5, 95% CI -5.12 to 0.12;  $P=.05$ ) (Table 2).

**Table 2.** Effect of the eHealth intervention based on the HIV risk prediction tool for men who have sex with men on HIV-related high-risk behaviors and intentions to undergo HIV testing (N=168).

HIV-related behavior	Intervention group (n=86)	Control group (n=82)	Effect size, OR <sup>a</sup> or mean difference <sup>b</sup> (95% CI)	P value
Proportion of participants having undergone HIV testing in the past 3 months, n (%)	75 (87)	68 (83)	1.30 <sup>a</sup> (0.55 to 3.09)	.55
Proportion of participants who intend to undergo HIV testing in the following 30 days, n (%)	77 (90)	65 (79)	2.20 <sup>a</sup> (0.90 to 5.35)	.07
Proportion of participants who used condoms in the past 3 months with casual sexual partners, n (%)	66 (87)	54 (70)	2.81 <sup>a</sup> (1.23 to 6.39)	.01
Proportion of participants who had passive anal intercourse in the past 3 months, n (%)	52 (61)	59 (72)	0.57 <sup>a</sup> (0.30 to 1.10)	.09
Proportion of participants who had unprotected passive anal intercourse in the past 3 months, n (%)	23 (27)	29 (35)	0.65 <sup>a</sup> (0.34 to 1.26)	.21
Proportion of participants engaging in group sex in the past 3 months, n (%)	6 (7)	7 (9)	0.80 <sup>a</sup> (0.26 to 2.50)	.14
Number of male sexual partners in the past 3 months, mean (SD)	3.5 (4.1)	6.0 (11.4)	-2.50 <sup>b</sup> (-5.12 to -0.12)	.05

<sup>a</sup>OR: odds ratio and 95% CI values have been used to indicate the effect size.

<sup>b</sup>Mean difference and 95% CI values have been used to indicate the effect size.

### Cost-Effectiveness Analysis

The total cost for the intervention group was US \$2577. The CER for the reduction in male sexual partners was US \$734.20,

and the ICER was US \$131.60. The CER for the promotion of condom usage with casual partners was US \$29.70, and the ICER was US \$19.70. Both ICERs were lower than the cost-effectiveness threshold for China (Table 3).



**Table 3.** Cost-effectiveness analysis of an eHealth intervention based on an HIV risk prediction model for men who have sex with men in China.

Groups	Effect	Cost (US \$)	Cost/effect (US \$)	Incremental cost-effectiveness ratio (US \$)
<b>Reduction in the number of sexual partners</b>				
Control, n	6.01	2248	374.0	N/A <sup>a</sup>
Intervention, n	3.51	2577	734.2	131.6
<b>Promotion of condom usage with casual sexual partners</b>				
Control, %	70	2248	32.1	N/A
Intervention, %	87	2577	29.7	19.7

<sup>a</sup>N/A: not applicable.

## Discussion

### Principal Findings

To our knowledge, this study is the first to assess the efficacy of an eHealth intervention based on an HIV risk prediction tool for the reduction of risk behavior and promotion of HIV testing among MSM and is a necessary step prior to the implementation of the predictive model in clinical practice. The number of male sexual partners in the intervention group significantly decreased during the study period, while the rate of insistence on condom usage among casual male sexual partners significantly increased. These findings indicate that eHealth interventions based on risk prediction might promote healthy sexual behavior among MSM.

This study found that a comprehensive online intervention based on risk assessment can significantly reduce the number of sexual partners among MSM. Previous studies have indicated that MSM with multiple sexual partners displayed continuous inconsistencies in terms of knowledge and behavior [17], probably owing to the insufficiency of conventional educational approaches to modify high-risk behaviors. In our study, the effectiveness of interventions in reducing the number of sexual partners may be attributed to an enhancement in risk perception through the personalized web-based risk evaluation component [18]. Therefore, the large-scale promotion and application of a personalized HIV risk prediction tool on social media platforms may increase risk perception and would be expected to reduce sexual network density among MSM.

Furthermore, we found that the risk prediction-based eHealth intervention could significantly increase condom usage with casual sexual partners among MSM. Pan et al [19] reported that MSM who seek sexual partners through web-based or social media platforms usually had more sexual partners, and notably more casual sexual partners, than those seeking sexual partners at other venues. Having unprotected intercourse with casual partners is an independent risk factor for HIV acquisition among MSM [19]. Thus, this intervention strategy aimed at improving HIV risk perception among MSM, and targeted interventions are useful for reducing the HIV risk among MSM who seek sexual partners on web-based or social media platforms [20]. In addition, the promotion of HIV testing is vital for preventing HIV infections, considering the low HIV testing rates in China. However, we observed only a marginal increase in the rate of participants intending to undergo HIV testing in the following 30 days in our study. One study performed in Peru [21] reported

that peer-mentored social media community-based interventions may improve the HIV testing rates among MSM. Therefore, the combination of social media community-based interventions with our risk prediction tool may increase the levels of HIV testing and reduce high-risk behavior in the future.

The clinical trial retention rate in our study approached 87.5%, which was comparable to a cluster randomized controlled trial on web-based peer education with a social media platform-based intervention to increase the HIV testing rate among MSM in Peru (90% retention rate at 12-week follow-up) [21]. This finding revealed that internet-based recruitment and the implementation of an eHealth intervention based on an HIV risk prediction tool are feasible. The higher clinical trial retention rates might be explained by the following reasons. First, monetary incentives in our study could improve the response enthusiasm to online surveys. Second, CBOs are familiar with MSM community members, understand their needs, and provide a trusted environment for communication among MSM. Generally, monetary incentives and CBOs could promote online recruitment and retention of MSM in internet-based studies aimed at improving service efficiency and effectiveness in preventing HIV infections [22,23]. Moreover, our results suggest that eHealth interventions based on risk prediction may help increase condom usage and decrease the number of male sexual partners among MSM. Compared with traditional facility-based HIV testing, our intervention could provide timely alerts to MSM with high-risk behaviors to make risk-reduction decisions and to provide information and resources on HIV prevention in a personalized manner [24]. Furthermore, the comprehensive online intervention model based on risk assessment used in this study was associated with a low cost. The cost-effectiveness of this intervention may be attributed to the conductance of this study on a social media platform, thus saving on housing costs and costs associated with other facilities. Additionally, an eHealth intervention can be extended to MSM who may not have time, resources, or the motivation to seek in-person preventive services.

### Limitations

This study has several limitations, which should be considered when extrapolating the results of our study. First, information collected from the questionnaires was self-reported, and laboratory data on sexually transmitted diseases (eg, HIV and syphilis) were not collected. Thus, we could not evaluate whether the intervention strategy had an influence on HIV or

sexually transmitted infections among MSM. Second, some MSM may not be frequent internet users, such as those with a low educational background and older MSM; therefore, they may not be well-suited for eHealth interventions. Therefore, other offline interventions should be developed as essential alternatives to web-based interventions to be extended to these groups. Third, the 12-week study duration may have contributed to a recall bias and telescoping errors for data collected through the web-based survey. Moreover, the extent to which MSM, who are familiar with local CBOs, who agree to participate in preventive interventions is different from MSM who are not familiar with local CBOs, thus resulting in a potential selection bias.

### Generalizability of the Findings

This study is applicable to MSM who use mobile phones and are willing to accept mobile phone-based interventions. The inclusion criterion of having a WeChat account limited the representativeness of the study population among the nationwide MSM population in China because MSM with a low educational background and older adult MSM may not be well-suited for mobile phone-based interventions. Therefore, our findings may not be generalizable to MSM not owning smartphones or to those who are not receptive to mobile phone-based interventions.

### Implications for Practice

These findings further the current understanding of the efficacy of comprehensive interventions based on HIV risk prediction models—which are delivered through social media platforms—on HIV-related behavioral changes among MSM, and provides a new paradigm for health interventions for MSM

and more opportunities for HIV surveillance and treatment, which have considerable implications and prospects.

### Future Prospects

Although our study demonstrates the efficacy of HIV risk prediction-based mobile phone interventions in promoting HIV testing and reducing high-risk behavior among MSM in China, the efficacy of this intervention in reducing the incidence of HIV or other sexually transmitted infections remains unclear owing to the lack of corresponding laboratory data and biological endpoints. Thus, future studies should collect laboratory data on HIV or other sexually transmitted infections and assess the efficacy of the intervention on epidemics of HIV or sexually transmitted infections among MSM. Furthermore, data on the proportion of MSM who are sex workers or MSM who have had female sexual partners should be collected in a future study to further the current understanding of the interaction of sexual networks among MSM and to verify the reliability of the survey data. Eventually, considering the important role of monetary incentives and CBOs in MSM recruitment and clinical trial retention, monetary incentives and mechanisms facilitating or supporting CBOs' engagement in effective and sustainable HIV/AIDS prevention programs should be considered in future peer studies.

### Conclusions

A mobile phone-based intervention based on an HIV risk prediction tool is feasible for MSM in China; this intervention could reduce the number of sexual partners and promote condom usage with casual sexual partners among MSM, thus providing a novel, convenient, and accessible intervention paradigm for this key population.

---

### Acknowledgments

We would like to thank Kang Qiang (leader, Shenyang Sunshine CBO) for his assistance in the recruitment and maintenance of the MSM in the clinical trial. This study was supported by the National Natural Science Foundation of China (81872674), the Mega-projects of National Science Research for the 13th Five-Year Plan (2017ZX10201101-002-007), the Central Public-interest Scientific Institution Basal Research Fund (2018PT31042), National Science and Technology Major Project (2018ZX10101-001-001-003), and Liaoning Natural Science Foundation Project (2020-BS-091).

---

### Authors' Contributions

JJX (xjjcnu@163.com) and HS (hongshang100@hotmail.com) both contribute equally as corresponding authors. KY, JJX, and HS designed the study. KY and JJX collected the data. KY and JJX analyzed the data, interpreted the results, and wrote the manuscript. All other authors provided their comments and critically revised the manuscript. All authors approved the manuscript before submission.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1127 KB - mhealth\\_v9i4e19511\\_app1.pdf](#)]

---

#### Multimedia Appendix 2

Checklist for Reporting Results of Internet E-Surveys.

[[DOCX File , 18 KB - mhealth\\_v9i4e19511\\_app2.docx](#)]

## Multimedia Appendix 3

Intervention index interface.

[\[PNG File , 341 KB - mhealth\\_v9i4e19511\\_app3.png \]](#)

## Multimedia Appendix 4

MSM risk assessment and tailored suggestions.

[\[PNG File , 396 KB - mhealth\\_v9i4e19511\\_app4.png \]](#)

## Multimedia Appendix 5

Application of free HIV self-testing kits and condoms.

[\[PNG File , 262 KB - mhealth\\_v9i4e19511\\_app5.png \]](#)

## Multimedia Appendix 6

Surrounding medical facility recommendation by geographical location.

[\[PNG File , 725 KB - mhealth\\_v9i4e19511\\_app6.png \]](#)

## Multimedia Appendix 7

HIV\_AIDS health education.

[\[PNG File , 638 KB - mhealth\\_v9i4e19511\\_app7.png \]](#)**References**

1. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet* 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60821-6](https://doi.org/10.1016/S0140-6736(12)60821-6)] [Medline: [22819660](https://pubmed.ncbi.nlm.nih.gov/22819660/)]
2. Xu J, Han X, Reilly KH, Shang H. New features of the HIV epidemic among men who have sex with men in China. *Emerg Microbes Infect* 2013 Jul;2(7):e45 [FREE Full text] [doi: [10.1038/emi.2013.45](https://doi.org/10.1038/emi.2013.45)] [Medline: [26038478](https://pubmed.ncbi.nlm.nih.gov/26038478/)]
3. Zhang W, Xu J, Zou H, Zhang J, Wang N, Shang H. HIV incidence and associated risk factors in men who have sex with men in Mainland China: an updated systematic review and meta-analysis. *Sex Health* 2016 May 26;13(4):373-382. [doi: [10.1071/SH16001](https://doi.org/10.1071/SH16001)] [Medline: [27294689](https://pubmed.ncbi.nlm.nih.gov/27294689/)]
4. Qin Q, Guo W, Tang W, Mahapatra T, Wang L, Zhang N, et al. Spatial Analysis of the Human Immunodeficiency Virus Epidemic among Men Who Have Sex with Men in China, 2006-2015. *Clin Infect Dis* 2017 Apr 01;64(7):956-963 [FREE Full text] [doi: [10.1093/cid/cix031](https://doi.org/10.1093/cid/cix031)] [Medline: [28362948](https://pubmed.ncbi.nlm.nih.gov/28362948/)]
5. Grosskopf NA, LeVasseur MT, Glaser DB. Use of the Internet and mobile-based "apps" for sex-seeking among men who have sex with men in New York City. *Am J Mens Health* 2014 Nov;8(6):510-520 [FREE Full text] [doi: [10.1177/1557988314527311](https://doi.org/10.1177/1557988314527311)] [Medline: [24658284](https://pubmed.ncbi.nlm.nih.gov/24658284/)]
6. Paz-Bailey G, Hoots BE, Xia M, Finlayson T, Prejean J, Purcell DW, NHBS Study Group. Trends in Internet Use Among Men Who Have Sex With Men in the United States. *J Acquir Immune Defic Syndr* 2017 Jul 01;75 Suppl 3:S288-S295 [FREE Full text] [doi: [10.1097/QAI.0000000000001404](https://doi.org/10.1097/QAI.0000000000001404)] [Medline: [28604430](https://pubmed.ncbi.nlm.nih.gov/28604430/)]
7. Wang Q, Ross MW. Differences between chat room and e-mail sampling approaches in Chinese men who have sex with men. *AIDS Educ Prev* 2002 Oct;14(5):361-366. [doi: [10.1521/aeap.14.6.361.24081](https://doi.org/10.1521/aeap.14.6.361.24081)] [Medline: [12413182](https://pubmed.ncbi.nlm.nih.gov/12413182/)]
8. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. *J Urban Health* 2013 Aug;90(4):729-739 [FREE Full text] [doi: [10.1007/s11524-012-9766-7](https://doi.org/10.1007/s11524-012-9766-7)] [Medline: [22983721](https://pubmed.ncbi.nlm.nih.gov/22983721/)]
9. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. *J Med Internet Res* 2014 May 26;16(5):e134 [FREE Full text] [doi: [10.2196/jmir.3393](https://doi.org/10.2196/jmir.3393)] [Medline: [24862459](https://pubmed.ncbi.nlm.nih.gov/24862459/)]
10. Hoenigl M, Chaillon A, Morris SR, Little SJ. HIV Infection Rates and Risk Behavior among Young Men undergoing community-based Testing in San Diego. *Sci Rep* 2016 May 16;6:25927 [FREE Full text] [doi: [10.1038/srep25927](https://doi.org/10.1038/srep25927)] [Medline: [27181715](https://pubmed.ncbi.nlm.nih.gov/27181715/)]
11. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. *AIDS Behav* 2014 Feb;18(2):285-296 [FREE Full text] [doi: [10.1007/s10461-013-0671-1](https://doi.org/10.1007/s10461-013-0671-1)] [Medline: [24292281](https://pubmed.ncbi.nlm.nih.gov/24292281/)]
12. Yun K, Xu J, Leuba S, Zhu Y, Zhang J, Chu Z, et al. Development and Validation of a Personalized Social Media Platform-Based HIV Incidence Risk Assessment Tool for Men Who Have Sex With Men in China. *J Med Internet Res* 2019 Jun 18;21(6):e13475 [FREE Full text] [doi: [10.2196/13475](https://doi.org/10.2196/13475)] [Medline: [31215509](https://pubmed.ncbi.nlm.nih.gov/31215509/)]
13. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]

14. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004 Sep 29;6(3):e34 [FREE Full text] [doi: [10.2196/jmir.6.3.e34](https://doi.org/10.2196/jmir.6.3.e34)] [Medline: [15471760](https://pubmed.ncbi.nlm.nih.gov/15471760/)]
15. Catania JA, Kegeles SM, Coates TJ. Towards an understanding of risk behavior: an AIDS risk reduction model (ARRM). *Health Educ Q* 1990;17(1):53-72. [doi: [10.1177/109019819001700107](https://doi.org/10.1177/109019819001700107)] [Medline: [2318652](https://pubmed.ncbi.nlm.nih.gov/2318652/)]
16. Cost effectiveness and strategic planning (WHO-CHOICE): Cost-effectiveness thresholds. World Health Organization. 2014. URL: <https://www.who.int/choice/cost-effectiveness/en/> [accessed 2020-07-01]
17. Liu XX, Zhang MH, Ye L, Li Z, Ruan Z, Cheng Z, et al. Study on influential factors of multiple sexual partner behaviors among MSM in Zhenjiang. *Jiangsu Journal of Preventive Medicine* 2016;27(4):411-413.
18. Ma G, Jiang Z, Chen H, Chen J. Influencing factors associated with multi-partner homosexual behaviors among MSM. *Anhui Journal of Preventive Medicine* 2013;19(6):414-416.
19. Pan S, Xu J, Han X, Zhang J, Hu Q, Chu Z, et al. Internet-Based Sex-Seeking Behavior Promotes HIV Infection Risk: A 6-Year Serial Cross-Sectional Survey to MSM in Shenyang, China. *Biomed Res Int* 2016;2016:2860346 [FREE Full text] [doi: [10.1155/2016/2860346](https://doi.org/10.1155/2016/2860346)] [Medline: [28105415](https://pubmed.ncbi.nlm.nih.gov/28105415/)]
20. Klein H. Condom use self-efficacy and HIV risk practices among men who use the internet to find male partners for unprotected sex. *Am J Mens Health* 2014 May;8(3):190-204 [FREE Full text] [doi: [10.1177/1557988313492172](https://doi.org/10.1177/1557988313492172)] [Medline: [23832954](https://pubmed.ncbi.nlm.nih.gov/23832954/)]
21. Young SD, Cumberland WG, Nianogo R, Menacho LA, Galea JT, Coates T. The HOPE social media intervention for global HIV prevention in Peru: a cluster randomised controlled trial. *Lancet HIV* 2015 Jan;2(1):e27-e32 [FREE Full text] [doi: [10.1016/S2352-3018\(14\)00006-X](https://doi.org/10.1016/S2352-3018(14)00006-X)] [Medline: [26236767](https://pubmed.ncbi.nlm.nih.gov/26236767/)]
22. Li H, Li JZ, Xiao Y, Sun JB, Wang JR, Ruan SM. The initiative process of AIDS high-risk interventions from global funding to governments purchasing services among men who have sex with men. *Chinese Journal of Disease Control* 2014;18(12):1162-1165.
23. Gaoxin Z, Hailing M, Junshi Z, Meizhu X, Yugang B, Jiming C. Exploration of government purchasing social organizational services in AIDS prevention and control. *Rural Health Management in China* 2015;35(1):53-55.
24. Wang S, Moss JR, Hiller JE. The cost-effectiveness of HIV voluntary counseling and testing in China. *Asia Pac J Public Health* 2011 Jul;23(4):620-633. [doi: [10.1177/1010539511412576](https://doi.org/10.1177/1010539511412576)] [Medline: [21727082](https://pubmed.ncbi.nlm.nih.gov/21727082/)]

## Abbreviations

- CBO:** community-based-organization  
**CER:** cost-effectiveness ratio  
**ICER:** incremental cost-effectiveness ratio  
**MSM:** men who have sex with men  
**OR:** odds ratio

*Edited by L Buis; submitted 24.04.20; peer-reviewed by M Wambura, C Hemingway, L Yongping; comments to author 12.06.20; revised version received 21.07.20; accepted 08.03.21; published 13.04.21.*

### *Please cite as:*

Yun K, Chu Z, Zhang J, Geng W, Jiang Y, Dong W, Shang H, Xu J  
*Mobile Phone Intervention Based on an HIV Risk Prediction Tool for HIV Prevention Among Men Who Have Sex With Men in China: Randomized Controlled Trial*  
*JMIR Mhealth Uhealth* 2021;9(4):e19511  
URL: <https://mhealth.jmir.org/2021/4/e19511>  
doi: [10.2196/19511](https://doi.org/10.2196/19511)  
PMID: [33847597](https://pubmed.ncbi.nlm.nih.gov/33847597/)

©Ke Yun, Zhenxing Chu, Jing Zhang, Wenqing Geng, Yongjun Jiang, Willa Dong, Hong Shang, Junjie Xu. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 13.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Feasibility and Preliminary Efficacy of a Community-Based Addiction Rehabilitation Electronic System in Substance Use Disorder: Pilot Randomized Controlled Trial

Xiaomin Xu<sup>1\*</sup>, MD; Shujuan Chen<sup>1\*</sup>, MD; Junning Chen<sup>2</sup>, BSc; Zhikang Chen<sup>1</sup>, BA; Liming Fu<sup>3</sup>, MSW; Dingchen Song<sup>4</sup>, BA; Min Zhao<sup>1,5,6,7</sup>, MD, PhD; Haifeng Jiang<sup>1,5</sup>, MD, PhD

<sup>1</sup>Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

<sup>2</sup>Nantong Winner Information Technology Co Ltd, Nantong, China

<sup>3</sup>Council of Shanghai Ziqiang Social Services, Shanghai, China

<sup>4</sup>Council of Shanghai Zhongzhi Social Services, Shanghai, China

<sup>5</sup>Shanghai Key Laboratory of Psychotic Disorders, Shanghai, China

<sup>6</sup>Shanghai Clinical Research Center for Mental Health, Shanghai, China

<sup>7</sup>Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

\*these authors contributed equally

**Corresponding Author:**

Haifeng Jiang, MD, PhD

Shanghai Mental Health Center

Shanghai Jiao Tong University School of Medicine

600 Wan Ping Nan Road

Shanghai,

China

Phone: 86 180 1731 1330

Email: [dragonjhf@hotmail.com](mailto:dragonjhf@hotmail.com)

## Abstract

**Background:** Drug use disorder has high potential for relapse and imposes an enormous burden on public health in China. Since the promulgation of the Anti-drug law in 2008, community-based rehabilitation has become the primary approach to treat drug addiction. However, multiple problems occurred in the implementation process, leading to a low detoxification rate in the community. Mobile health (mHealth) serves as a promising tool to improve the effectiveness and efficiency of community-based rehabilitation. Community-based addiction rehabilitation electronic system (CAREs) is an interactive system for drug users and their assigned social workers.

**Objective:** The study aimed to examine the feasibility and preliminary efficacy of CAREs in community-based rehabilitation from the perspective of drug users and social workers in Shanghai, China.

**Methods:** In this pilot randomized controlled trial, 40 participants were recruited from the community in Shanghai from January to May 2019. Participants randomized to the intervention group (n=20) received CAREs + community-based rehabilitation, while participants in the control group (n=20) received community-based rehabilitation only for 6 months. CAREs provided education, assessment, and SOS (support) functions for drug users. The assigned social workers provided service and monitored drug use behavior as usual except that the social workers in the intervention group could access the webpage end to obtain drug users' information and fit their routine workflow into CAREs. The primary outcome was the feasibility of CAREs, reflected in the overall proportion and frequency of CAREs features used in both app and webpage end. The secondary outcomes were the effectiveness of CAREs, including the percentage of drug-positive samples, longest period of abstinence, contact times with social workers, and the change of Addiction Severity Index (ASI) from baseline to the 6-month follow-up.

**Results:** The number of participants logged in to the app ranged from 7 to 20 per week, and CAREs had relatively high levels of continued patient use. Drug users preferred assessment and education features in the app end while their social workers showed high levels of use in urine results record and viewing assessment results on the webpage end. After the 6-month intervention, 3.3% (17/520) of samples in the intervention group and 7.5% (39/520) in the control group were drug-positive ( $F=4.358$ ,  $P=.04$ ).

No significant differences were noted between the control and intervention groups in terms of longest duration of abstinence, number of contact times and ASI composite scores.

**Conclusions:** The study preliminarily demonstrated that with relatively good feasibility and acceptability, CAREs may improve the effectiveness and efficiency of the community-based rehabilitation, which provided instruction for further improvement of the system.

**Trial Registration:** ClinicalTrials.gov NCT03451344; <https://clinicaltrials.gov/ct2/show/NCT03451344>

**International Registered Report Identifier (IRRID):** RR2-10.3389/fpsy.2018.00556

(*JMIR Mhealth Uhealth* 2021;9(4):e21087) doi:[10.2196/21087](https://doi.org/10.2196/21087)

## KEYWORDS

mobile health; drug use; rehabilitation; community health service; China

## Introduction

In China, illicit drug abuse is an increasingly serious and complicated problem. An estimated 2.4 million people used illicit drugs by the end of 2018, imposing a significant burden on the addiction treatment system [1]. According to articles 33, 38, and 49 of the Anti-Drug Law, drug users are ordered to receive treatment for addiction in three forms: medical institution-based detoxification, community-based rehabilitation, and isolated compulsory treatment [2], among which the community-based rehabilitation is the primary approach for drug rehabilitation in China [3].

Community-based rehabilitation provides counseling and support, monitors drug use behavior to help drug users maintain long-term abstinence and promote their social integration. As the most economically developed city in China, Shanghai has more than 1000 social workers who serve as not only the main supervisor but also the helper for drug users, playing a significant role in the implementation of community-based rehabilitation [4]. However, due to scarcity of effective prevention and intervention within the community, drug users cannot mitigate risks associated with drug use, which may result in adverse consequences: relapse, incarceration, or death [5]. This dilemma arises in part from the immature development of community-based rehabilitation, which is reflected in inadequate facilities and funding sources in community or nonprofessional services and delayed feedback from social workers [6]. It is also a challenge for both social workers and drug users to maintain 3-year drug rehabilitation according to national legislation [2]. Therefore, it is essential to develop novel interventions to improve the effectiveness and efficiency of community-based rehabilitation, which can provide enormous benefit to public health.

Bringing mobile health (mHealth) into the routine treatment regimen may make it possible by delivering evidence-based health information, ongoing monitoring, and personalized intervention according to collected data on patients via sensors, apps, webpages, and location-tracking technology, which can improve treatment adherence, patient-provider communication, and recovery from diseases [7,8]. The past decade has seen the emergence of mHealth for chronic disease management including substance abuse [9], primarily in the means of short message service (SMS, or text messaging) or phone calls [10]. Compared with traditional technologies, smartphones show

significant advantages for supporting complicated apps, accessing measurements with built-in mobile sensors, and allowing an omnipresent internet connection [11]. In China, smartphones are now widely used, with around 713 million users in 2018 [12]. Strategies for drug rehabilitation such as information or education, social support, assessment, feedback, monitoring, skills training, psychological intervention, self-management, and relaxation could be realized by smartphones app [13,14]. Although several apps for substance-related and addictive disorders have proved to be effective in randomized clinical trials, most of the apps are designed for alcohol and nicotine abuse [15,16]. Meanwhile, to the best of our knowledge, no app interventions with integrated functions for improving drug users' antirelapse skills, increasing working efficiency, and enhancing interaction between social workers and drug users have been used in community-based rehabilitation.

Due to the current context in China, our research team developed a community-based addiction rehabilitation electronic system (CAREs) centered on a smartphone app with the aim of improving the professionalization and efficiency of the community-based rehabilitation [17]. This paper reports findings from a pilot randomized controlled trial demonstrating the feasibility and preliminary efficacy of CAREs in community-based rehabilitation.

## Methods

### Study Design

This study was a randomized controlled trial with 2 parallel groups comparing preliminary efficacy of CAREs + community-based rehabilitation to community-based rehabilitation alone. The protocol was registered at ClinicalTrials.gov [NCT03451344] and previously published [17]. The study was approved by the Shanghai Mental Health Center Ethics Committee (2017-33) and was in accordance with the principles of the Declaration of Helsinki.

### Participants

Participants newly designated to receive community-based rehabilitation were enrolled (January 2019 to May 2019) from the social worker station in Shanghai in this study. All participants were recruited through advertisements in the Ziqiang and Zhongzhi consortia, the 2 largest specialized social worker consortia in Shanghai to help drug users in the community.

Inclusion criteria were aged 20 to 50 years, met the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria for substance use disorders (SUD), and provided informed consent. Exclusion criteria were inability or refusal to use smartphone app, severe cognitive impairment, or a history of suicidality.

## Procedures

After signing the written informed consent, participants were screened for eligibility within 3 days. At this time, the demographic characteristics, drug use information, and urine drug screen (UDS) were collected by trained social workers. Participants who met the inclusion criteria received a 7-day training with their assigned social workers on how to use CAREs to familiarize them uniformly with it. Participants were randomly assigned to receive either CAREs + community-based rehabilitation or community-based rehabilitation only (1:1 ratio) for 6 months using simple randomization tables generated by SPSS Statistics version 22 (IBM Corp).

Each participant was assessed for severity of problems associated with drugs in 7 domains (alcohol use, drug use, medical, employment, legal, family/social, and psychiatric status) using the Chinese version of the Addiction Severity Index (ASI) at baseline and after the 6-month intervention. The Chinese version of the ASI has been proven with good reliability and validity [18-20]. UDS was collected once a week during the study course (26 weeks in total); samples were screened for heroin, amphetamine-type stimulants, marijuana, cocaine, and ketamine.

Social workers conducted assessments and weekly UDS in the social worker station of each subdistrict. The people who provided the CAREs intervention were different from those who performed UDS and assessment of ASI. However, it was not possible to completely blind the evaluators to group allocation as the participants in intervention group may talk about the CAREs intervention during the assessment. Participants in the intervention group were compensated with 50 RMB (US \$7.64) for potential cost for mobile data, and at the end of the study they received a smartphone with the CAREs app installed.

## Intervention

### **Control Group: Community-Based Rehabilitation**

According to national legislation [2], drug users who received regular community-based rehabilitation must sign an agreement to comply when they were newly enrolled in this program. They submitted a written report if they left their city of residence, in accordance with the localized management. Participants visited their assigned social worker and agreed to be tested for illicit drugs every 2 months. In this study, participants in the control and intervention groups were asked to submit a urine sample once a week. As those whose urine test result was positive would be sent to an isolated compulsory treatment center for 2 years, it was conservatively estimated that participants who did not or refused to submit urine samples were considered to have positive results on UDS. At other times, social workers helped their clients apply for social benefits as needed and provided counseling irregularly if necessary.

### **Intervention Group: CAREs + Community-Based Rehabilitation**

The intervention group received the same community-based rehabilitation as the control group. In addition, they received access to CAREs, which consists of a smartphone app for drug users and a webpage for social workers. Participants were required to log in to the app at least once a week and encouraged to use it at other times. The app was designed with 3 major modules for drug users providing education, assessment, and coping skills and support.

The educational content was selected from a course on saying no to drugs and delivered in the form of text or video with material covering basic knowledge about drugs, confidence building, treatment principles, antidrug skills, and emotion management using cognitive behavioral therapy based on the relapse prevention model [21]. More specifically, clinical guidelines for stimulant-induced mental or somatic symptoms, replacement therapy, methadone maintenance treatment, the role of personal relationships in addiction treatment, and so on were introduced in the treatment principles component, and antidrug skills contained an introduction of self-help groups to create a support network, develop clear thinking about major life events and stress, resist temptation from all sides, cope with cravings, etc. These educational resources were delivered in a specific order once a week. Users were expected to complete the educational courses on time when receiving the reminder. To reinforce learning and withdrawal motivation, participants were encouraged to revisit lessons at any time and could earn points by viewing educational content and finishing the corresponding exercises after learning.

Assessments were provided using 5 self-report instruments with proven reliability in substance-using samples. Craving was assessed by visual analog scale [22], with 0 cm being no craving at all and 10 cm suggesting the most craving ever experienced. Scores of the Patient Health Questionnaire-9 [23] indicated the level of depression severity: minimal (0-4), mild (5-9), moderate (10-14), moderately severe (15-19), or severe (20-27). Scores of the General Anxiety Disorder-7 [24] indicated the level of anxiety severity: minimal (0-4), mild (5-9), moderate (10-14), or severe (15-21). The Alcohol Use Disorders Identification Test [25] was used to screen at-risk drinking: low-risk alcohol use (0-7), hazardous alcohol use (8-15), harmful alcohol use (16-19), or alcohol dependence ( $\geq 20$ ). The family and employment status was adapted from the Chinese version of the ASI [18-20] and contained 6 yes/no questions, with 0-3 points indicating a nonideal family or employment status. Users received scores to indicate level of severity as well as the real-time feedback accordingly. If the scores were above the normal cutoff point, users were advised to learn the coping skills from the education and SOS module or turn to their assigned social workers. Meanwhile, the social workers would receive reminders to pay more attention to those users and initiate safety protocols when necessary. Participants were required to complete the assessment once a week.

The SOS module contained tools for skills coping, relaxation, and call forwarding to connect with contacts, including family members, doctors, and social workers. Skills coping with craving

are vital due to the close connection between craving and relapse [26]. Relaxation training included music relaxation and abdominal breathing. Drug users could also interact with their assigned social workers on the message board through the app.

Meanwhile, social workers could obtain all information in CAREs apps of their assigned drug users through the webpage end. By using the webpage, social workers could record urine test results, track real-time location, review results of assessments, reply to messages, and send reminders if participants missed a urine test or assessment. If drug users left the supervision area in Shanghai without notifying the social worker, the system automatically initiated alarm but only the matched social worker received the reminder. Of note, all data would be stored in the server for safety and privacy considerations.

### Outcome

The primary outcome was feasibility of CAREs in the community-based rehabilitation program, reflected in the use of CAREs in both app and webpage ends, reported as the overall proportion and frequency of CAREs features used. Use was defined as a participant or social worker accessing a feature page (not the home page) in CAREs. However, data on mean count per user was not possible for CAREs due to technical reasons on the database end.

Secondary outcomes were effectiveness of CAREs, including (1) UDS results examined in overall percentage of drug-positive samples, (2) longest duration of sustained abstinence, defined as the greatest number of consecutive weeks of negative UDS samples in the 6-month period, (3) ASI composite scores

summed according to the answers to each of the problem areas, and (4) contact times, days when participants interacted with their assigned social workers in the form of sending messages via CAREs, SMS, face-to-face meeting, or phone call.

### Statistical Analysis

Intention-to-treat principle was used in all analyses. The statistical analyses were conducted with SPSS Statistics version 22. Significance level was set at  $P < .05$ . Descriptive statistics were used to describe baseline data and CAREs feature use. Chi-square test for categorical variables and Student  $t$  test for continuous variables were used to examine the baseline comparability of the two groups. Percentage of drug-positive samples, longest period of abstinence, and contact times were examined for significant difference by analyses of variance. As for the ASI composite scores, repeated measure analysis of variance was used to evaluate the differences between baseline and 6-month scores. G\*power program [27] was used for power analysis.

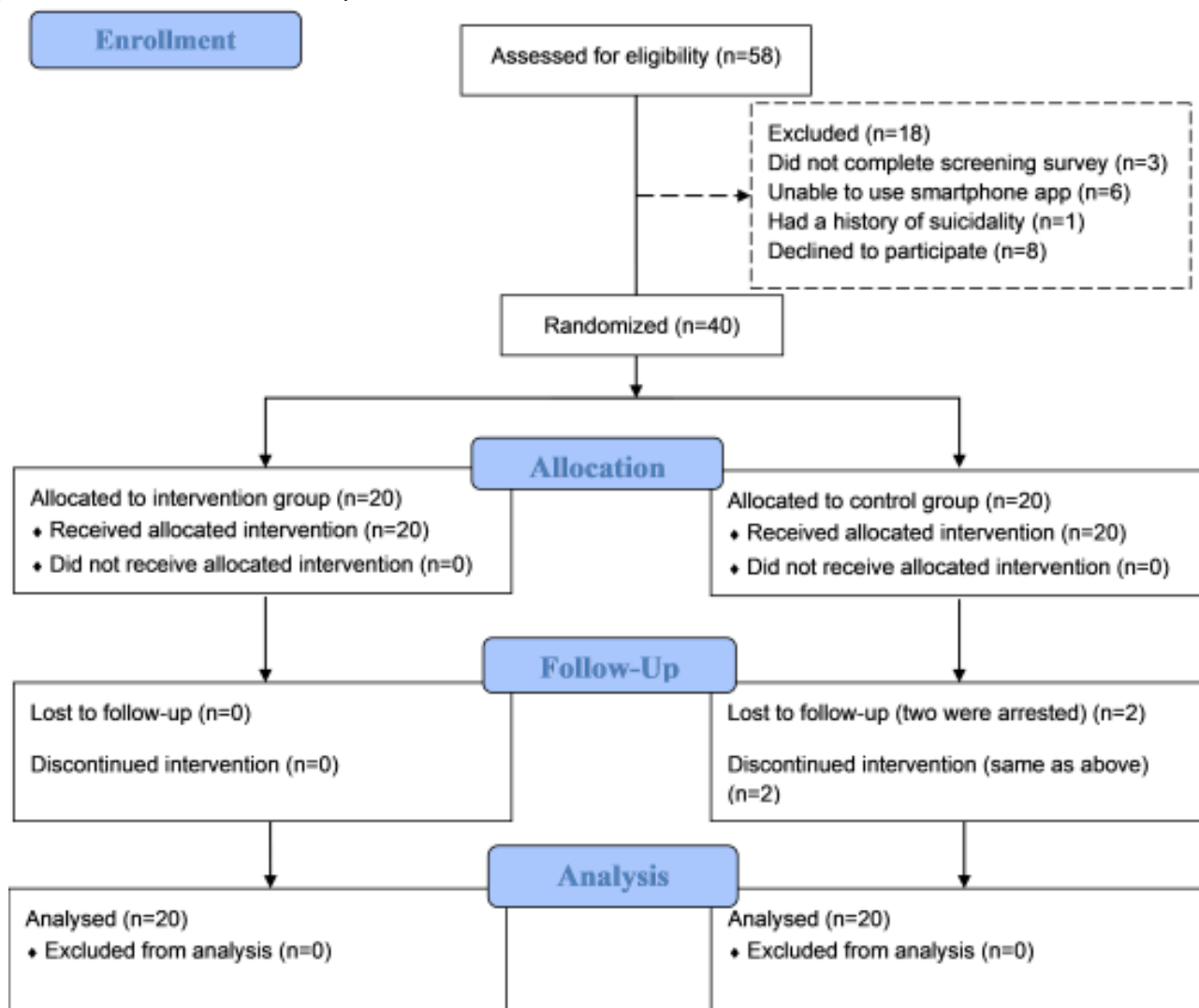
## Results

### Participant Characteristics

Figure 1 shows the participant flow. A total of 40 people were randomized into the intervention ( $n=20$ ) and control ( $n=20$ ) group of the trial, with only 2 participants (5%) in the control group dropping out due to being arrested (caught using illicit drugs by the police at weeks 20 and 25, respectively). Baseline demographic characteristics were presented in Table 1. No significant differences between the groups were found ( $P > .05$ ).



Figure 1. CONSORT flowchart of the study.



**Table 1.** Baseline and demographic characteristics of participants (n=40).

Characteristic	Intervention (n=20)	Control (n=20)	$t/\chi^2$	<i>P</i> value
Age in years, mean (SD)	47.0 (8.8)	45.1 (11.0)	0.604	.55
Male, n (%)	16 (80)	15 (75)	0.143	.71
Employed, n (%)	9 (45)	6 (30)	0.960	.33
<b>Marital status, n (%)</b>	— <sup>a</sup>	—	0.895	.24
Married	6 (30)	11 (55)	—	—
Divorced	8 (40)	4 (20)	—	—
Never	6 (30)	5 (25)	—	—
<b>Education, n (%)</b>	—	—	1.423	.49
Low (<6 years)	0 (0)	1 (5)	—	—
Middle (6-9 years)	14 (70)	13 (65)	—	—
High (>9 years)	6 (30)	6 (30)	—	—
Accumulated years of drug use, mean (SD)	10.7 (7.9)	10.1 (7.2)	0.272	.79
<b>Type of primary drug use, n (%)</b>	—	—	0.360	.55
Methamphetamine	19 (95)	18 (90)	—	—
Heroin	1 (5)	2 (10)	—	—

<sup>a</sup>Not applicable.

## Use of CAREs

### Features on Webpage End for Social Workers

Uses of CAREs features on both webpage and app ends were shown in Table 2. For social workers, the 3 most commonly used functions were recording urine test results, location tracking, and viewing assessment information of drug users, with 100% (20/20), 80% (16/20), and 60% (12/20), respectively, of social workers using them on at least one occasion. Reminders for urine test and assessment were the least used. However, when it came to mean number of times over the whole study

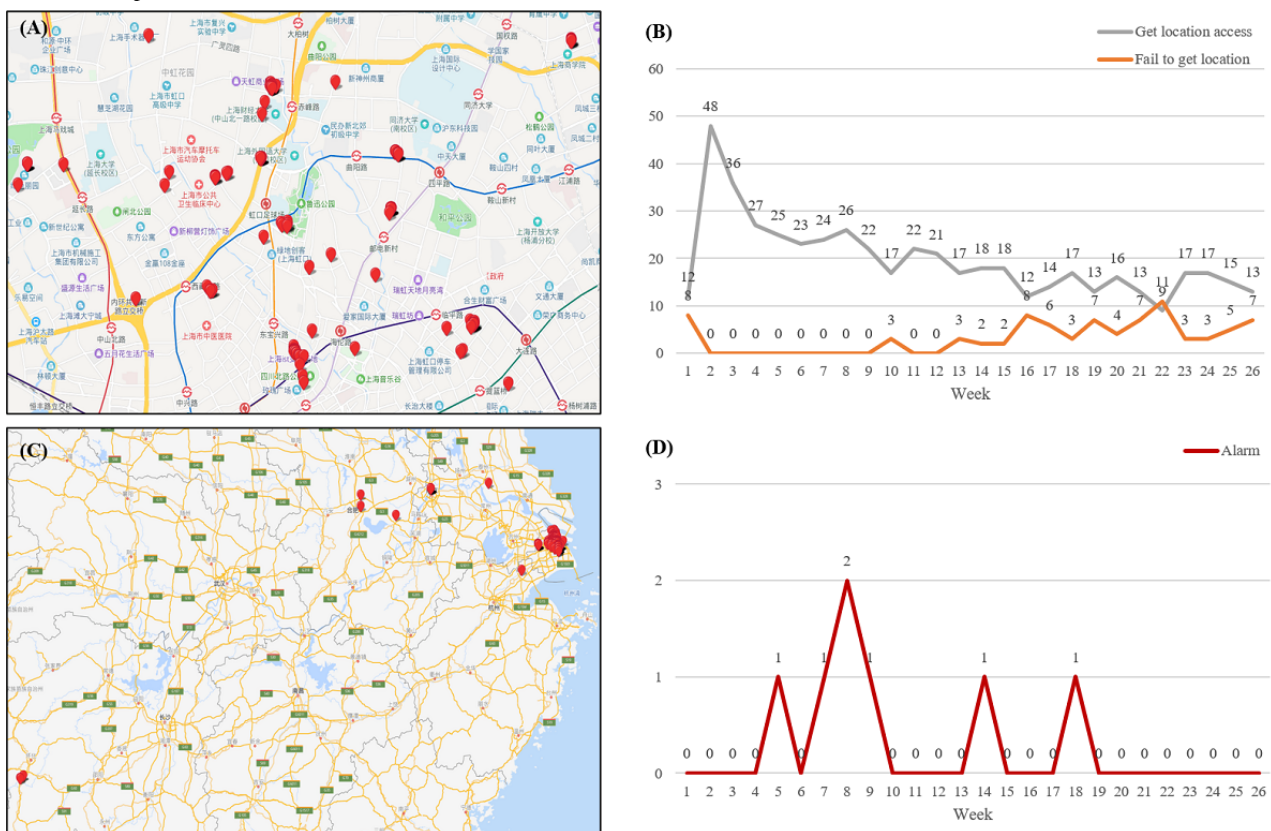
course, use of these features was relatively low. Over the study period, social workers accessed urine tests record and assessment results an average of 32.87 times and 26.56 times, respectively. Location tracking was the third most frequently used, and data from the back-end showed that the number of times accessing location of participants in the intervention group decreased over time, while the system failed to get some locations in the latter half of the study (Figure 2A and B). When the drug users left the supervision area in Shanghai without submitting a written report, the system automatically initiated alarm 7 times in total (Figure 2C and D).

**Table 2.** Use of CAREs features on both webpage and app ends.

Features	Rate of users with at least one use (%)	Mean total <sup>a</sup>
<b>Social workers (webpage end)</b>		
Record urine test results	20 (100)	32.87
Location tracking	16 (80)	12.04
View assessment information	12 (60)	26.56
Send urine test reminder	7 (35)	3.47
Send assessment reminder	3 (15)	0.65
<b>Drug users (app end)</b>		
Assessment	20 (100)	24.15
Education	15 (75)	63.3
<b>SOS</b>		
Music relaxation	9 (45)	2.65
Abdominal breathing	7 (35)	1.6
Hotline for doctors	3 (15)	0.25
Hotline for voluntary drug rehabilitation center	3 (15)	0.25
Call family number	9 (45)	1.75

<sup>a</sup>Mean total: mean number of times across the participants in the intervention group over the whole study course.

**Figure 2.** Location-tracking feature for social workers to monitor participants in intervention group: (A) close shot of Shanghai: drug users moving around within the supervision area; (B) number of locations accessed and not accessed (failed) from CAREs app per week; (C) remote view of Shanghai: some drug users had left the supervision area without reporting to the matched social workers, and the system automatically initiated alarm; and (D) number of alarms per week.



**Features on App End for Drug Users**

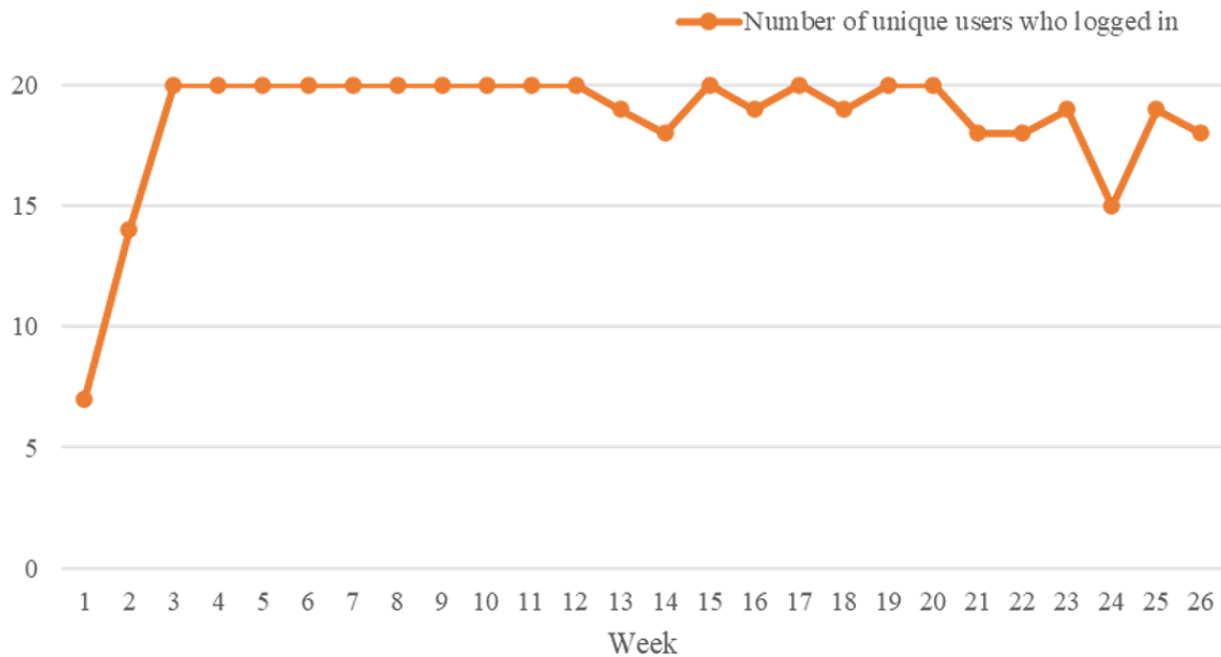
The number of participants logged in to the app ranged from 7 to 20 per week, and CAREs had high levels of continued patient

use, as shown in Figure 3. In terms of content analysis of CAREs app for drug users (Table 2), the assessment feature was accessed by the largest number of users (20/20, 100%), followed

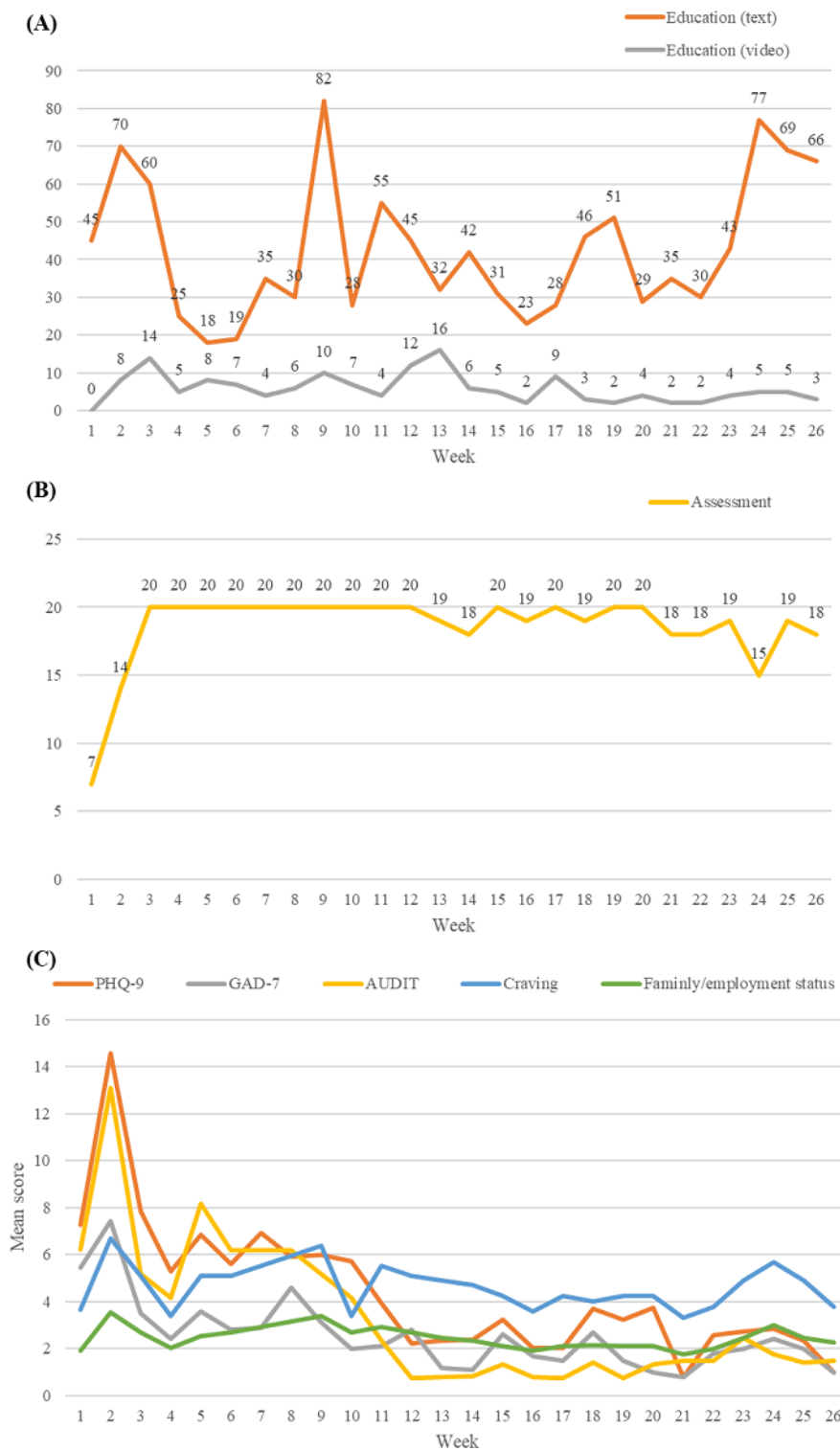
by education (15/20, 75%). However, the greatest average number of uses over the 6-month period was education, which was shown in Figure 4A. Educational content delivered by text message was preferred by participants over video. Number of unique users who used the assessment function of the CAREs app per week is shown in Figure 4B, and Figure 4C displays the assessment results per week. At the beginning of the study,

especially in the second week, mean scores of assessments indicated that participants were generally in the moderately severe range of depression, mild anxiety, and hazardous drinking. These mean scores decreased over time and maintained a low level in the last few weeks. However, drug users maintained nonideal family and employment status and the same level of craving during the 6-month period.

**Figure 3.** Number of unique participants who log in to the CAREs app at least once each week.



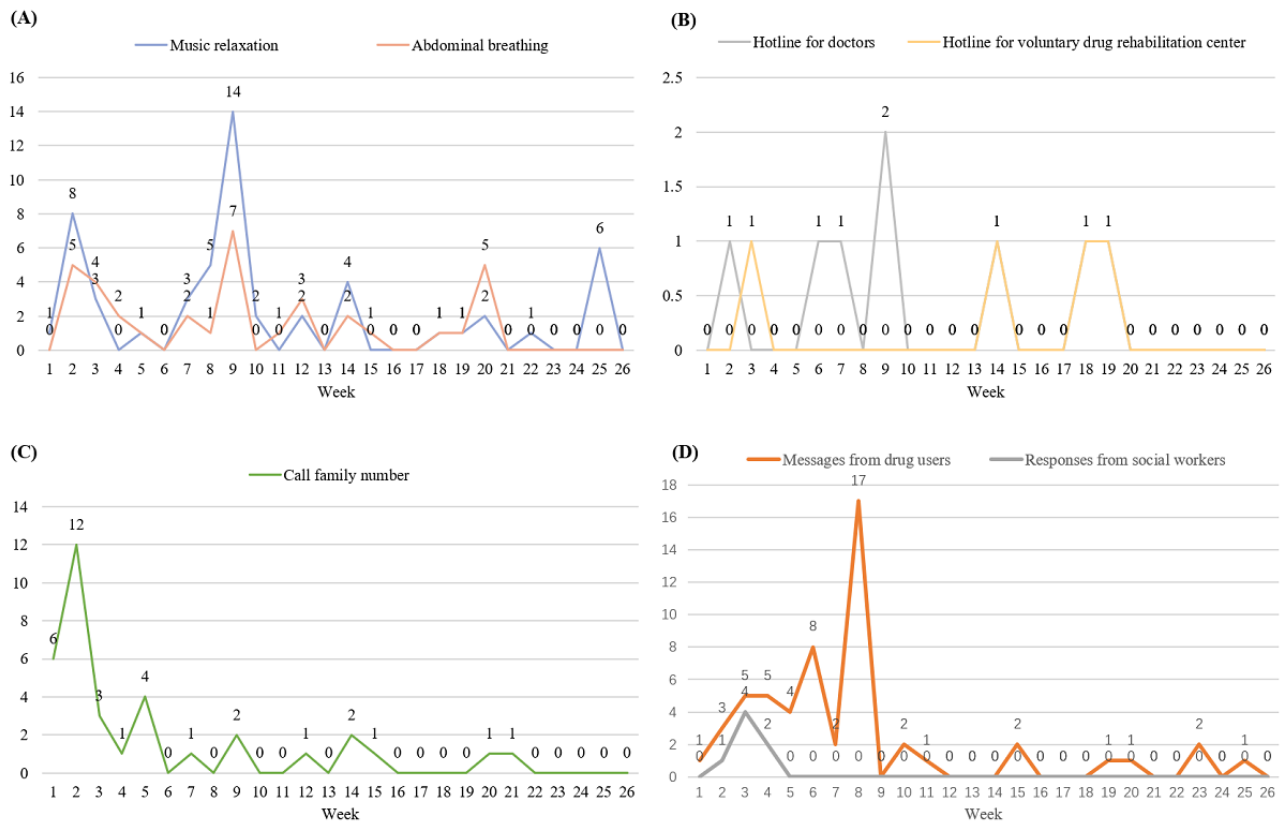
**Figure 4.** (A) Total number of people using education (in the form of text or video) function from CAREs app per week; (B) number of unique users who used assessment function from CAREs app per week (maximum number of people was 20); (C) mean scores of assessment results decreased over time.



Data on SOS functions showed low levels of use (Table 2), with a relatively small proportion of participants choosing music relaxation or abdominal breathing to cope with craving and emergency. Figure 5A to 5C showed that the frequency of use of some SOS features (such as calling family members) peaked at first but decreased over time, while the hotline for doctors

and voluntary drug rehabilitation center maintained low frequency of use all the time. As for the message board shown in Figure 5D, drug users left messages frequently in the first few weeks but decreased at seventh week, with low responses from social workers all the time.

**Figure 5.** Mean number of times of SOS features per week: (A) relaxation training (including music relaxation and abdominal breathing); (B) call forwarding service to doctors and voluntary drug rehabilitation center; (C) call forwarding service to family members; (D) message board used in both drug user and social worker ends.



**Drug-Related Outcomes and Contact Times**

The follow-up assessment was conducted 0 days to 5 days postintervention (mean 0.4), and there was no difference between groups ( $t=11.427, P=.06$ ). Descriptive and statistical tests are shown in Table 3. There were 987 urine samples collected in total during the 6-month study period (weeks 20 to 26; mean 24.7). Participants in the intervention group showed

a lower percentage of drug-positive samples than participants in the control group. No significant differences were found in terms of longest period of abstinence, ASI composite scores, or contact times. The power to compare percentage of drug-positive samples between the 2 groups was 0.74, with longest period of abstinence 0.73, contact times 0.16, and ASI composite scores ranging from 0.07 to 0.75.

**Table 3.** Effectiveness of outcomes: drug use and contact times.

Variables	Intervention (n=20)		Control (n=20)		F score	P value
	Baseline	End of 6 months	Baseline	End of 6 months		
% drug-positive samples, mean (SD)	— <sup>a</sup>	3.3 (5.0)	—	7.5 (7.5)	4.358	.04
Longest period of abstinence, mean (SD)	—	24.65 (2.21)	—	22.80 (3.59)	3.851	.06
Contact times, mean (SD)	—	42.45 (18.51)	—	36.65 (17.48)	1.038	.32
<b>ASI<sup>b</sup> composite scores, mean (SD)</b>						
Medical status	0.21 (0.27)	0.07 (0.29)	0.15 (0.26)	0.04 (0.13)	0.316 <sup>c</sup>	.58
Employment status	0.74 (0.28)	0.07 (0.22)	0.67 (0.31)	0.05 (0.23)	0.636 <sup>c</sup>	.43
Alcohol use	0.05 (0.80)	0.01 (0.03)	0.13 (0.16)	0.02 (0.15)	1.672 <sup>c</sup>	.20
Drug use	0.05 (0.07)	0.01 (0.03)	0.04 (0.04)	0.01 (0.03)	0.302 <sup>c</sup>	.59
Legal status	0.04 (0.08)	0.00 (0.11)	0.05 (0.13)	0.02 (0.08)	0.025 <sup>c</sup>	.87
Family/social status	0.10 (0.13)	0.03 (0.10)	0.10 (0.13)	0.04 (0.11)	0.011 <sup>c</sup>	.92
Psychiatric status	0.06 (0.06)	0.03 (0.06)	0.05 (0.07)	0.03 (0.07)	0.043 <sup>c</sup>	.84

<sup>a</sup>Not applicable.

<sup>b</sup>ASI: Addiction Severity Index.

<sup>c</sup>Statistical tests represent effects of group × time interaction.

## Discussion

### Principal Findings

To our knowledge, this was the first pilot study examining the feasibility and preliminary efficacy of a smartphone app (CAREs) to facilitate community-based rehabilitation program from the perspective of the drug users and social workers in China. The findings of this study provided preliminary evidence of CAREs as a potential tool with moderately good acceptability and effectiveness among individuals mainly using methamphetamine from community-based rehabilitation program. This was reflected in better performance on the UDS in the intervention group and moderate to high engagement with some CAREs features. Unexpectedly, the previous hypothesis that CAREs intervention would significantly prolong longest period of sustained abstinence, increase contact times, and reduce ASI composite scores was not confirmed in this study.

Growing awareness of the impact of addiction on public health calls for broader reach with lower barriers to services for drug users [28,29]. However, traditional support services may bring drug users concerns about relevance and stigma, creating barriers to access to routine treatment [30,31]. Under the addiction treatment model and system in contemporary China, studies examining the feasibility of mHealth among drug users have mainly been conducted in patients using heroin from methadone maintenance treatment (MMT) programs. Han [32] found poor acceptance of a mobile phone-based ecological momentary assessment app in the MMT population. Liang [4], however, demonstrated the feasibility and potential benefits for participants in the MMT program to receive both surveys and text messages from the S-Health app. As there is still a large deficiency in social workers' time and professional competence to provide evidence-based interventions in community-based

rehabilitation program [33], the significance of the role of social workers was also considered in this study. CAREs is an interactive system for drug users and their matched social workers to address problems of addiction.

Like other mHealth app interventions [34,35], although nearly all participants logged in to CAREs during each week in the 6-month period, a gradual decline of app engagement was found in this study. Some flaws of CAREs may hinder operations. For example, some clients could not log in to their accounts or view repeated content because of technical difficulties. Nevertheless, the functions of self-monitoring, education, urine results records, and viewing assessment results showed relatively high use. As shown in this study, long-term education and gradually improved assessment results may contribute to the lower percentage of drug-positive samples in the intervention group. Some participants in the intervention group mentioned that they felt proud when the assessment results improved and became more confident to overcome the addiction by knowing more about drug-related knowledge. According to the social cognitive theory [36,37], self-monitoring such as the assessment and education features of the CAREs app would increase self-efficacy beliefs which operate with goals, positive outcome expectancies, and environmental perception to facilitate one's motivation and behaviors. In line with previous studies of SUD intervention, repeated assessments improved self-monitoring [38,39]. Education increased awareness of potential risks [40-42] and skills to prevent relapse [43,44] and thereby helped individuals change their dysfunctional behaviors.

On the other hand, compared with the existing routine face-to-face interviews, the help and education of social workers was more convenient through the trial implementation of CAREs, especially in terms of real-time feedback. Social workers requested that the CAREs app data be integrated into

their existing management system, which could fit into their existing workflow, improve their work efficiency, and help them be more familiar with drug users' situations. Some social workers mentioned that negative affect of drug users in the intervention group was detected earlier and more easily through the assessment information, and they would pay more attention to that drug user accordingly.

Although use of the location tracking with automatic alarm function was moderate, it indeed provided an effective method for better supervision of drug users within the city of residence. However, some drug users turned these features off with privacy concerns.

Additionally, use of SOS functions was relatively high initially but dropped to a low level of use in the later of the study course. Users may have accessed the features at first only because of freshness. The reason for low use of the call forwarding service and message board probably was that individuals with SUD preferred face-to-face interviews, which provided a way for them to communicate with others [32]. Similar finding has been reported in a feature analysis of a smartphone-based smoking cessation app: few participants used the Phone a Counselor function, as weekly offline counseling sessions made the hotline unnecessary or redundant [45]. Therefore, contact times between drug users and their assigned social workers in the intervention group was more than that in the control group, but not significantly.

In terms of reduced ASI composite scores, this was probably because ASI covers several domains associated with drugs, and CAREs did not provide components for legal and employment domains. This suggests that future iterations of the CAREs app should integrate more comprehensive components with the help of multicollaborations such as specialty addiction treatment settings, communities, and related administrative departments.

### Limitations

This study had several limitations. First, it should be acknowledged that this study was conducted only in Shanghai,

where social workers were provided who specialized in helping drug users in the community due to legal requirements; furthermore, CAREs was designed with social workers as service providers. Therefore, considering different antidrug systems and procedures in other regions of China, the results of this study are for reference only. Second, the study terminated with only 40 samples because of the emergence of COVID-19, and the majority of the subjects were males. Although this study is promising, generalizability is limited by the small sample sizes and gender difference. Third, the mean count per user data could not be obtained from the database for technical reasons. As a pilot study, the relatively small sample size limited statistical power to detect the efficacy of CAREs, so the paper predominantly focused on the app's feasibility. Technical support for the CAREs database should be improved to conduct secondary analysis of improved behavioral outcomes associated with the use of CAREs with larger samples. Also, technical problems may have prevented some users from engaging with the app. The related technical problems should be recorded in a future research process to distinguish whether problematic engagement is because of unwillingness or inability to use. Fourth, data on previous drug-related history (eg, UDS, period of abstinence, and treatment patterns) and information associated with comorbid symptoms were not collected in this study. More information should be requested in future studies to exclude the potential bias. Fifth, as addiction is a chronic disease [46], follow-up data are essential to evaluate whether the treatment is effective in maintaining abstinent after a 6-month intervention.

### Conclusions

This pilot study suggests a moderate level of feasibility and acceptability for CAREs in a community-based rehabilitation program. It preliminarily demonstrated that the support offered by CAREs may improve the effectiveness and efficiency of community-based rehabilitation. Future studies will focus on updating CAREs and conducting long-term effectiveness trials in well-powered and larger samples to improve the quality of rehabilitation for drug addiction in China.

---

### Acknowledgments

This work was supported by grant DLY201818 from the Clinical Research Center, Shanghai Jiao Tong University School of Medicine; grant 2017YQ013 from the Municipal Human Resources Development Program for Outstanding Young Talents in Medical and Health Sciences in Shanghai; grant 18411961200 from the Science and Technology Innovation Plan in Shanghai; grant 13DZ2260500 from the Shanghai Key Laboratory of Psychotic Disorders; grant 2017YFC1310400 from the National Key R&D Program of China; grant 81771436 from National Nature Science Foundation; and grant 19MC1911100 from the Shanghai Clinical Research Center for Mental Health.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 2378 KB - mhealth\\_v9i4e21087\\_app1.pdf](#)]

---

### References



1. 2018 Annual Report on Drug Control in China. China Nation Narcotic Control Commission. URL: [http://www.nncc626.com/2019-06/17/c\\_1210161797.htm](http://www.nncc626.com/2019-06/17/c_1210161797.htm) [accessed 2021-03-24]
2. Anti-Drug Law of the People's Republic of China. URL: [http://www.china.org.cn/china/LegislationsForm2001-2010/2011-02/11/content\\_21899159.htm](http://www.china.org.cn/china/LegislationsForm2001-2010/2011-02/11/content_21899159.htm) [accessed 2021-03-24]
3. Luo T, Wang J, Li Y, Wang X, Tan L, Deng Q, et al. Stigmatization of people with drug dependence in China: a community-based study in Hunan province. *Drug Alcohol Depend* 2014 Jan 01;134:285-289. [doi: [10.1016/j.drugalcdep.2013.10.015](https://doi.org/10.1016/j.drugalcdep.2013.10.015)] [Medline: [24239068](https://pubmed.ncbi.nlm.nih.gov/24239068/)]
4. Liang D, Han H, Du J, Zhao M, Hser Y. A pilot study of a smartphone application supporting recovery from drug addiction. *J Subst Abuse Treat* 2018 May;88:51-58 [FREE Full text] [doi: [10.1016/j.jsat.2018.02.006](https://doi.org/10.1016/j.jsat.2018.02.006)] [Medline: [29606226](https://pubmed.ncbi.nlm.nih.gov/29606226/)]
5. Haifeng J, Di L, Jiang D, Haiming S, Zhikang C, Liming F, et al. Gender differences in recovery consequences among heroin dependent patients after compulsory treatment programs. *Sci Rep* 2015 Dec 08;5:17974 [FREE Full text] [doi: [10.1038/srep17974](https://doi.org/10.1038/srep17974)] [Medline: [26644283](https://pubmed.ncbi.nlm.nih.gov/26644283/)]
6. Zhuang H, Zhang P. Comparative Survey on Community Drug Abandonment and Rehabilitation at Home and Abroad. *J Yunnan Police Off Acad* 2010;1:60-63. [doi: [10.3969/j.issn.1672-6057.2010.01.012](https://doi.org/10.3969/j.issn.1672-6057.2010.01.012)]
7. Wang Y, Min J, Khuri J, Xue H, Xie B, A Kaminsky L, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2020 Apr 28;8(4):e15400 [FREE Full text] [doi: [10.2196/15400](https://doi.org/10.2196/15400)] [Medline: [32343253](https://pubmed.ncbi.nlm.nih.gov/32343253/)]
8. Sim I. Mobile Devices and Health. *N Engl J Med* 2019 Sep 05;381(10):956-968. [doi: [10.1056/NEJMra1806949](https://doi.org/10.1056/NEJMra1806949)] [Medline: [31483966](https://pubmed.ncbi.nlm.nih.gov/31483966/)]
9. Berrouguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [FREE Full text] [doi: [10.2196/jmir.5066](https://doi.org/10.2196/jmir.5066)] [Medline: [27287668](https://pubmed.ncbi.nlm.nih.gov/27287668/)]
10. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: integrating smartphone technology in behavioral healthcare. *Prof Psychol Res Pract* 2011 Dec;42(6):505-512. [doi: [10.1037/a0024485](https://doi.org/10.1037/a0024485)]
11. Miralles I, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone apps for the treatment of mental disorders: systematic review. *JMIR Mhealth Uhealth* 2020 Apr 02;8(4):e14897 [FREE Full text] [doi: [10.2196/14897](https://doi.org/10.2196/14897)] [Medline: [32238332](https://pubmed.ncbi.nlm.nih.gov/32238332/)]
12. [Number of smartphone users in China from 2017 to 2023]. URL: <https://www.iimedia.cn/c1061/67391.html> [accessed 2021-03-24]
13. Chapman C, Champion KE, Birrell L, Deen H, Brierley M, Stapinski LA, et al. Smartphone apps about crystal methamphetamine ("ice"): systematic search in app stores and assessment of composition and quality. *JMIR Mhealth Uhealth* 2018 Nov 21;6(11):e10442 [FREE Full text] [doi: [10.2196/10442](https://doi.org/10.2196/10442)] [Medline: [30463834](https://pubmed.ncbi.nlm.nih.gov/30463834/)]
14. Meredith SE, Alessi SM, Petry NM. Smartphone applications to reduce alcohol consumption and help patients with alcohol use disorder: a state-of-the-art review. *Adv Health Care Technol* 2015;1:47-54 [FREE Full text] [doi: [10.2147/AHCT.S65791](https://doi.org/10.2147/AHCT.S65791)] [Medline: [27478863](https://pubmed.ncbi.nlm.nih.gov/27478863/)]
15. Kruse CS, Lee K, Watson JB, Lobo LG, Stoppelmoor AG, Oyibo SE. Measures of effectiveness, efficiency, and quality of telemedicine in the management of alcohol abuse, addiction, and rehabilitation: systematic review. *J Med Internet Res* 2020 Jan 31;22(1):e13252 [FREE Full text] [doi: [10.2196/13252](https://doi.org/10.2196/13252)] [Medline: [32012048](https://pubmed.ncbi.nlm.nih.gov/32012048/)]
16. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. *Transl Behav Med* 2017 Jun;7(2):292-299 [FREE Full text] [doi: [10.1007/s13142-017-0492-2](https://doi.org/10.1007/s13142-017-0492-2)] [Medline: [28527027](https://pubmed.ncbi.nlm.nih.gov/28527027/)]
17. Wang Z, Chen S, Chen J, Xu C, Chen Z, Zhuang W, et al. A community-based addiction rehabilitation electronic system to improve treatment outcomes in drug abusers: protocol for a randomized controlled trial. *Front Psychiatry* 2018;9:556 [FREE Full text] [doi: [10.3389/fpsy.2018.00556](https://doi.org/10.3389/fpsy.2018.00556)] [Medline: [30459653](https://pubmed.ncbi.nlm.nih.gov/30459653/)]
18. Luo W, Wu Z, Wei X, Jia W, Zhang Q, Li L. Adaptation and evaluation of Chinese version of addiction severity index-V and its usage for assessing addiction of patients in methadone maintenance treatment clinics. *Chin J Drug Depend* 2007;16(3):373-378. [doi: [10.3969/j.issn.1007-9718.2007.05.014](https://doi.org/10.3969/j.issn.1007-9718.2007.05.014)]
19. Zhao M, Li X, Hao W, Wang Z, Zhang M, Xu D. A preliminary study of the reliability and validity of the Addiction Severity Index. *J Tradit Chin Med* 2004;4:679-680. [doi: [10.1111/j.1360-0443.2003.00665.x](https://doi.org/10.1111/j.1360-0443.2003.00665.x)]
20. Sun Z, Chen H, Su Z, Zhou X, Zhang S, Hao W, et al. The Chinese version of the Addiction Severity Index (ASI-C): reliability, validity, and responsiveness in Chinese patients with alcohol dependence. *Alcohol* 2012 Dec;46(8):777-781. [doi: [10.1016/j.alcohol.2012.08.005](https://doi.org/10.1016/j.alcohol.2012.08.005)] [Medline: [23146314](https://pubmed.ncbi.nlm.nih.gov/23146314/)]
21. Larimer ME, Palmer RS, Marlatt GA. Relapse prevention. An overview of Marlatt's cognitive-behavioral model. *Alcohol Res Health* 1999;23(2):151-160 [FREE Full text] [Medline: [10890810](https://pubmed.ncbi.nlm.nih.gov/10890810/)]
22. Kleykamp BA, De Santis M, Dworkin RH, Huhn AS, Kampman KM, Montoya ID, et al. Craving and opioid use disorder: a scoping review. *Drug Alcohol Depend* 2019 Dec 01;205:107639. [doi: [10.1016/j.drugalcdep.2019.107639](https://doi.org/10.1016/j.drugalcdep.2019.107639)] [Medline: [31683241](https://pubmed.ncbi.nlm.nih.gov/31683241/)]

23. Delgadillo J, Payne S, Gilbody S, Godfrey C, Gore S, Jessop D, et al. How reliable is depression screening in alcohol and drug users? A validation of brief and ultra-brief questionnaires. *J Affect Disord* 2011 Nov;134(1-3):266-271. [doi: [10.1016/j.jad.2011.06.017](https://doi.org/10.1016/j.jad.2011.06.017)] [Medline: [21723619](https://pubmed.ncbi.nlm.nih.gov/21723619/)]
24. Delgadillo J, Payne S, Gilbody S, Godfrey C, Gore S, Jessop D, et al. Brief case finding tools for anxiety disorders: validation of GAD-7 and GAD-2 in addictions treatment. *Drug Alcohol Depend* 2012 Sep 01;125(1-2):37-42. [doi: [10.1016/j.drugalcdep.2012.03.011](https://doi.org/10.1016/j.drugalcdep.2012.03.011)] [Medline: [22480667](https://pubmed.ncbi.nlm.nih.gov/22480667/)]
25. Zhang C, Yang GP, Li Z, Li XN, Li Y, Hu J, et al. [Reliability and validity of the Chinese version on Alcohol Use Disorders Identification Test]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2017 Aug 10;38(8):1064-1067. [doi: [10.3760/cma.j.issn.0254-6450.2017.08.013](https://doi.org/10.3760/cma.j.issn.0254-6450.2017.08.013)] [Medline: [28847055](https://pubmed.ncbi.nlm.nih.gov/28847055/)]
26. Serre F, Fatseas M, Swendsen J, Auriacombe M. Ecological momentary assessment in the investigation of craving and substance use in daily life: a systematic review. *Drug Alcohol Depend* 2015 Mar 1;148:1-20. [doi: [10.1016/j.drugalcdep.2014.12.024](https://doi.org/10.1016/j.drugalcdep.2014.12.024)] [Medline: [25637078](https://pubmed.ncbi.nlm.nih.gov/25637078/)]
27. Erdfelder E, Faul F, Buchner A. GPOWER: a general power analysis program. *Behav Res Methods Instr Comput* 1996 Mar;28(1):1-11. [doi: [10.3758/bf03203630](https://doi.org/10.3758/bf03203630)]
28. Campbell ANC, Nunes EV, Matthews AG, Stitzer M, Miele GM, Polsky D, et al. Internet-delivered treatment for substance abuse: a multisite randomized controlled trial. *Am J Psychiatry* 2014 Jun;171(6):683-690 [FREE Full text] [doi: [10.1176/appi.ajp.2014.13081055](https://doi.org/10.1176/appi.ajp.2014.13081055)] [Medline: [24700332](https://pubmed.ncbi.nlm.nih.gov/24700332/)]
29. Buck JA. The looming expansion and transformation of public substance abuse treatment under the Affordable Care Act. *Health Aff (Millwood)* 2011 Aug;30(8):1402-1410. [doi: [10.1377/hlthaff.2011.0480](https://doi.org/10.1377/hlthaff.2011.0480)] [Medline: [21821557](https://pubmed.ncbi.nlm.nih.gov/21821557/)]
30. Tait RJ, McKetin R, Kay-Lambkin F, Carron-Arthur B, Bennett A, Bennett K, et al. A web-based intervention for users of amphetamine-type stimulants: 3-month outcomes of a randomized controlled trial. *JMIR Ment Health* 2014;1(1):e1 [FREE Full text] [doi: [10.2196/mental.3278](https://doi.org/10.2196/mental.3278)] [Medline: [26543901](https://pubmed.ncbi.nlm.nih.gov/26543901/)]
31. Chan RCH, Mak WWS. Common sense model of mental illness: understanding the impact of cognitive and emotional representations of mental illness on recovery through the mediation of self-stigma. *Psychiatry Res* 2016 Dec 30;246:16-24. [doi: [10.1016/j.psychres.2016.09.013](https://doi.org/10.1016/j.psychres.2016.09.013)] [Medline: [27644016](https://pubmed.ncbi.nlm.nih.gov/27644016/)]
32. Han H, Zhang JY, Hser Y, Liang D, Li X, Wang SS, et al. Feasibility of a mobile phone app to support recovery from addiction in china: secondary analysis of a pilot study. *JMIR Mhealth Uhealth* 2018 Feb 27;6(2):e46 [FREE Full text] [doi: [10.2196/mhealth.8388](https://doi.org/10.2196/mhealth.8388)] [Medline: [29487040](https://pubmed.ncbi.nlm.nih.gov/29487040/)]
33. McLellan AT, Carise D, Kleber HD. Can the national addiction treatment infrastructure support the public's demand for quality care? *J Subst Abuse Treat* 2003 Sep;25(2):117-121. [Medline: [14680015](https://pubmed.ncbi.nlm.nih.gov/14680015/)]
34. Brown J, Michie S, Geraghty AWA, Yardley L, Gardner B, Shahab L, et al. Internet-based intervention for smoking cessation (StopAdvisor) in people with low and high socioeconomic status: a randomised controlled trial. *Lancet Respir Med* 2014 Dec;2(12):997-1006 [FREE Full text] [doi: [10.1016/S2213-2600\(14\)70195-X](https://doi.org/10.1016/S2213-2600(14)70195-X)] [Medline: [25262458](https://pubmed.ncbi.nlm.nih.gov/25262458/)]
35. Taylor GMJ, Dalili MN, Semwal M, Civiljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2017 Sep 04;9:CD007078. [doi: [10.1002/14651858.CD007078.pub5](https://doi.org/10.1002/14651858.CD007078.pub5)] [Medline: [28869775](https://pubmed.ncbi.nlm.nih.gov/28869775/)]
36. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004 Apr;31(2):143-164. [doi: [10.1177/1090198104263660](https://doi.org/10.1177/1090198104263660)] [Medline: [15090118](https://pubmed.ncbi.nlm.nih.gov/15090118/)]
37. Bandura A. Health promotion from the perspective of social cognitive theory. *Psychol Health* 1998 Jul;13(4):623-649. [doi: [10.1080/08870449808407422](https://doi.org/10.1080/08870449808407422)]
38. Jenkins RJ, McAlaney J, McCambridge J. Change over time in alcohol consumption in control groups in brief intervention studies: systematic review and meta-regression study. *Drug Alcohol Depend* 2009 Feb 01;100(1-2):107-114. [doi: [10.1016/j.drugalcdep.2008.09.016](https://doi.org/10.1016/j.drugalcdep.2008.09.016)] [Medline: [19041196](https://pubmed.ncbi.nlm.nih.gov/19041196/)]
39. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. *J Med Internet Res* 2017 Mar 24;19(3):e90 [FREE Full text] [doi: [10.2196/jmir.7111](https://doi.org/10.2196/jmir.7111)] [Medline: [28341617](https://pubmed.ncbi.nlm.nih.gov/28341617/)]
40. Carrà G, Crocamo C, Humphris G, Tabacchi T, Bartoli F, Neufeind J, et al. Engagement in the Overdose RiSk InforMationN (ORION) e-Health tool for opioid overdose prevention and self-efficacy: a preliminary study. *Cyberpsychol Behav Soc Netw* 2017 Dec;20(12):762-768. [doi: [10.1089/cyber.2016.0744](https://doi.org/10.1089/cyber.2016.0744)] [Medline: [29148826](https://pubmed.ncbi.nlm.nih.gov/29148826/)]
41. Darke S, Mills KL, Ross J, Teesson M. Rates and correlates of mortality amongst heroin users: findings from the Australian Treatment Outcome Study (ATOS), 2001-2009. *Drug Alcohol Depend* 2011 Jun 01;115(3):190-195. [doi: [10.1016/j.drugalcdep.2010.10.021](https://doi.org/10.1016/j.drugalcdep.2010.10.021)] [Medline: [21130585](https://pubmed.ncbi.nlm.nih.gov/21130585/)]
42. Ochalek TA, Heil SH, Higgins ST, Badger GJ, Sigmon SC. A novel mHealth application for improving HIV and Hepatitis C knowledge in individuals with opioid use disorder: a pilot study. *Drug Alcohol Depend* 2018 Sep 01;190:224-228 [FREE Full text] [doi: [10.1016/j.drugalcdep.2018.05.032](https://doi.org/10.1016/j.drugalcdep.2018.05.032)] [Medline: [30056321](https://pubmed.ncbi.nlm.nih.gov/30056321/)]
43. Takano A, Miyamoto Y, Kawakami N, Matsumoto T. Web-based cognitive behavioral relapse prevention program with tailored feedback for people with methamphetamine and other drug use problems: development and usability study. *JMIR Ment Health* 2016 Jan;3(1):e1 [FREE Full text] [doi: [10.2196/mental.4875](https://doi.org/10.2196/mental.4875)] [Medline: [26740264](https://pubmed.ncbi.nlm.nih.gov/26740264/)]

44. Goldstein SP, Evans BC, Flack D, Juarascio A, Manasse S, Zhang F, et al. Return of the JITAI: applying a just-in-time adaptive intervention framework to the development of m-health solutions for addictive behaviors. *Int J Behav Med* 2017 Oct;24(5):673-682 [FREE Full text] [doi: [10.1007/s12529-016-9627-y](https://doi.org/10.1007/s12529-016-9627-y)] [Medline: [28083725](https://pubmed.ncbi.nlm.nih.gov/28083725/)]
45. Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. *J Med Internet Res* 2016 Dec 12;18(12):e321 [FREE Full text] [doi: [10.2196/jmir.6058](https://doi.org/10.2196/jmir.6058)] [Medline: [27956375](https://pubmed.ncbi.nlm.nih.gov/27956375/)]
46. Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry* 2016 Aug;3(8):760-773 [FREE Full text] [doi: [10.1016/S2215-0366\(16\)00104-8](https://doi.org/10.1016/S2215-0366(16)00104-8)] [Medline: [27475769](https://pubmed.ncbi.nlm.nih.gov/27475769/)]

## Abbreviations

**ASI:** Addiction Severity Index

**CAREs:** community-based addiction rehabilitation electronic system

**mHealth:** mobile health

**MMT:** methadone maintenance treatment

**SMS:** short message service

**SUD:** substance use disorders

**UDS:** urine drug screen

*Edited by L Buis; submitted 05.06.20; peer-reviewed by K Prior, J Tung; comments to author 10.07.20; revised version received 03.08.20; accepted 22.03.21; published 16.04.21.*

*Please cite as:*

*Xu X, Chen S, Chen J, Chen Z, Fu L, Song D, Zhao M, Jiang H*

*Feasibility and Preliminary Efficacy of a Community-Based Addiction Rehabilitation Electronic System in Substance Use Disorder: Pilot Randomized Controlled Trial*

*JMIR Mhealth Uhealth* 2021;9(4):e21087

URL: <https://mhealth.jmir.org/2021/4/e21087>

doi: [10.2196/21087](https://doi.org/10.2196/21087)

PMID: [33861211](https://pubmed.ncbi.nlm.nih.gov/33861211/)

©Xiaomin Xu, Shujuan Chen, Junning Chen, Zhikang Chen, Liming Fu, Dingchen Song, Min Zhao, Haifeng Jiang. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 16.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Development of the Niggle App for Supporting Young People on Their Dynamic Journey to Well-being: Co-design and Qualitative Research Study

Stoyan R Stoyanov<sup>1,2</sup>, MRes; Oksana Zelenko<sup>1</sup>, PhD; Aleksandra Staneva<sup>1,3</sup>, PhD, MRes; David J Kavanagh<sup>4</sup>, PhD; Calvin Smith<sup>5</sup>, PhD; Gavin Sade<sup>6</sup>, PhD; Jessica Cheers<sup>1</sup>, BA Fine Arts (Hons); Leanne Hides<sup>7</sup>, PhD

<sup>1</sup>QUT Design Lab, Creative Industries Faculty, Queensland University of Technology, Brisbane, Australia

<sup>2</sup>Division of Advocacy and Research, yourtown, Brisbane, Australia

<sup>3</sup>HORÓ Counseling Centre for Women, Brisbane, Australia

<sup>4</sup>Centre for Children's Health Research, School of Psychology and Counselling, Queensland University of Technology, Brisbane, Australia

<sup>5</sup>Office of Medical Education, Faculty of Medicine, The University of Queensland, Brisbane, Australia

<sup>6</sup>Creative Industries Faculty, Queensland University of Technology, Brisbane, Australia

<sup>7</sup>School of Psychology, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Australia

**Corresponding Author:**

Stoyan R Stoyanov, MRes

QUT Design Lab

Creative Industries Faculty

Queensland University of Technology

2 George Street

Brisbane, 4000

Australia

Phone: 61 31380316

Email: [stoyan.stoyanov@qut.edu.au](mailto:stoyan.stoyanov@qut.edu.au)

## Abstract

**Background:** Adolescence is a life stage characterized by intense development and increased vulnerability. Yet, young people rarely seek help for mental health, often due to stigma and embarrassment. Alarming, even those who do seek help may not be able to receive it. Interventions focused on well-being offer a protective factor against adversity. Highly effective, innovative, theoretically sound, accessible, and engaging mobile health (mHealth) interventions that can be used to look beyond mental ill-health and toward mental well-being are urgently needed.

**Objective:** We aimed to explore how young Australians conceptualize and construct recovery journeys from feeling unwell to being well in order to inform the conceptual design of a youth-led information-, resource-, and support-focused mHealth intervention.

**Methods:** A sample of young people, grouped by age (12-15 years, 16-19 years, and 20-25 years), took part in 3 in-person participatory design workshops (per group). Young people's understanding and representation of well-being, feeling unwell, and the recovery journey were investigated using visual and linguistic data collection methods: photo elicitation and journey mapping. A social constructionist perspective was used for thematic analysis to produce a conceptual model of the recovery journey. A mobile app was co-designed and all app functions were mapped through iterative development and testing by young people and a team of psychology, research, design and information technology experts.

**Results:** Young people (n=25) described a 6-stage journey with specific barriers and coping strategies. The findings, when situated within the personal recovery framework in mental health, emphasize the cyclic and iterative model of change. Through co-design, the new app—Niggle—was conceptualized as a visual representation of an amorphous problem, which can be addressed through app functions corresponding to the most helpful strategies that young people used to progress through the stages of their recovery journey.

**Conclusions:** Niggle is available to offer support to young people for a range of problems and provides a hot link to counseling services in Australia. This paper elaborates on the process of in-depth qualitative data collection through visual, linguistic, and

co-design methods. The findings of this study give insight into young people's understanding of well-being and recovery. This paper could aid the development of high-quality personalized mHealth interventions and support resources.

(*JMIR Mhealth Uhealth* 2021;9(4):e21085) doi:[10.2196/21085](https://doi.org/10.2196/21085)

## KEYWORDS

mHealth; adolescence; youth; young people; well-being; co-design; participatory design; qualitative research; thematic analysis; recovery; visual methods

## Introduction

### Young People and Mental Health

Adolescence is characterized by intense cognitive, emotional, social, and physiological growth; increased vulnerability [1,2]; and deficits in emotion regulation capability [3-6]. Young people (under the age of 25 years) experience high levels of mental disorders [1,7] and increased exposure to risk factors [8,9]. In recent years, however, there has been a shift of focus in psychology from mental illness and pathologizing ill health toward positive psychology approaches, well-being, recovery-oriented approaches, and resilience- and strength-based frameworks, which is exemplified in research on well-being.

### Well-being

Commonly described as the experience of positive emotions about one's life, such as happiness, life satisfaction, and positive functioning, well-being is illustrated by a sense of fulfilment and engagement in life [10]. In 1998, Keyes [11] proposed the dual-continua mental health model which posits that mental illness and mental well-being lie on two separate but related continua, such that mental well-being cannot be merely defined as the absence of mental illness. Further research suggested that mental illness and well-being represent two distinct subcomponents of an overarching construct of mental health [12]. Well-being recovery models describe the transcendence of symptoms into a renewed sense of meaningful life, despite the limitations of mental illness [13]. Greater well-being is associated with fewer symptoms of mental illness and reduced incidences of behavioral issues such as criminality, substance use [14,15], and physical illness [16]. Thus, recent approaches to mental health propose a shift from the treatment and prevention of mental illness toward the enhancement of well-being [17].

### Lack of Well-being in Young Australians

Keyes [18] defines *flourishing* as the presence of emotional, psychological, and social well-being. The lack of well-being, defined as *languishing*, is characterized by low levels of these characteristics. Young Australians consistently report lower levels of well-being than adults [9,19-22]. According to these definitions, lower youth well-being levels reflect emotional, psychological, social, and help-seeking-related factors. (1) Emotional—Young adulthood is a developmental period frequently associated with an increased exposure to risk factors including increased emotional imbalance, developing emotion regulation skills, and the stressful period of transition to adulthood, placing demands on coping resources [8,9]. Young people experience increasing emotional unrest and have low capacity to effectively regulate emotions [4,5]. (2)

Psychological—Most mental illness emerges before the age of 25 years [7]. In Australia, mental disorders commonly affect children and adolescents (aged 4 to 17 years), with 13.9% reporting a mental disorder in the past 12 months [23]. The 2019 National Health Survey in Australia [24] revealed that young people aged 15 to 24 years experience the highest levels of mental or behavioral conditions compared to all other age groups. (3) Social—Relationships play a central role in the well-being of young people [25]. The most common reason for individuals aged 12 to 25 years contacting major Australian youth counseling services such as Headspace or Kids Helpline was “relationship problems,” which includes family, partner, or peer relationships [26,27]. (4) Help Seeking—It is concerning that only one-third of young people with mental disorders seek help. This is due to barriers such as stigma, confidentiality issues, lack of accessibility, self-reliance, lack of knowledge about mental health services, and fear or stress about the act of help seeking or the help source itself [28-31]. Alarming, even those who seek help may struggle to receive it. For example, Kids Helpline, Australia's largest telephone and web-counseling service, is unable to meet demand, with continuously decreasing yearly response rates of approximately 50% [32,33]. Highly effective, innovative, accessible, and engaging support solutions that promote well-being are urgently needed [29].

### The Promise of Mobile Health Technology

The ubiquity and increasing functionality of modern mobile health (mHealth) technology offers promise for filling the gap by addressing barriers, such as stigma, confidentiality concerns, or difficulty of accessing support, to help seeking [34]. Young people are among the highest adopters of innovative technologies [34]. This presents an excellent opportunity for the development of novel, accessible mHealth interventions that offer reliable information and support.

Quality mHealth interventions are engaging, functional, professionally designed, and contain reliable information [35,36]. Several factors are essential to increase an intervention's potential for uptake, use and efficacy. This includes involving stakeholders in co-design, through participatory design workshops; conducting in-depth research of the phenomena which the mHealth intervention is designed to address [12,37,38]; involving researchers and design experts to interpret and translate participatory design workshop data into the design of the intervention; and involving health experts to assess its alignment with established theoretical models [39].

The purpose of this project was to qualitatively explore how young people conceptualize and construct individual recovery journeys, from being unwell toward experiencing greater levels of well-being. The findings were used to inform the conceptual

design of a youth-led information-, resource-, and support-focused mHealth intervention for the largest support service for young people in Australia (Kids Helpline).

## Methods

### Participant Recruitment

Ethics approval from Queensland University of Technology was granted for all stages of this project (Human Research Ethics Approval Number 1600000956). Recruitment was done via Facebook and researcher networks within Brisbane, Australia. Young people were invited to attend a series of workshops with the aim of co-designing a new digital tool for youth well-being. Quota sampling was applied and included age (12-25 years) and English language as participation criteria. Information sheets were sent to 35 young people who contacted the team. We grouped participants by age (3 separate age groups were formed—12-15 years; 16-19 years, and 20-25 years) to ensure variability of age-specific experiences and to reduce age-related peer pressure within each group. Consent (including parental consent for the youngest cohort) was obtained before participation. A series of participatory design workshops (3 per age group) were conducted between February and March 2017; each workshop lasted approximately 2 hours each. Participants were offered \$30 (approximately US \$23.25) or a movie voucher for their time.

### Research Design

The participatory design workshops [38,40,41] combined visual and linguistic data collection methods—photo elicitation [42,43] and journey mapping—drawn from user experience design [44,45], to ensure a rich variety of concepts. The use of linguistic methods alone creates limitations, especially for younger participants [42,46-48]; therefore, the young people were provided with multiple communication modalities to explore individual experiences and understanding [45,49,50] of well-being, being unwell, the recovery journey between, and potential technological solutions.

### Procedure and Analysis

#### Step 1

To immerse participants in the topic of well-being without priming them with researchers' definitions, we first asked them

to create lists of associations for the terms *well-being* and *being unwell*.

#### Step 2

Each participant was offered a set of 126 images cut out from different sources (Multimedia Appendix 1), which included a large variety of colorful or black-and-white images, patterns, textures, and icons that were representative of relationships, emotions, nature, activities, popular app icons, and abstract images. The images in the set were aggregated and refined by the research team and a group of youth consultants. Participants were invited to (1) select images associated with well-being, (2) select images associated with being unwell, and (3) create separate collages for each category.

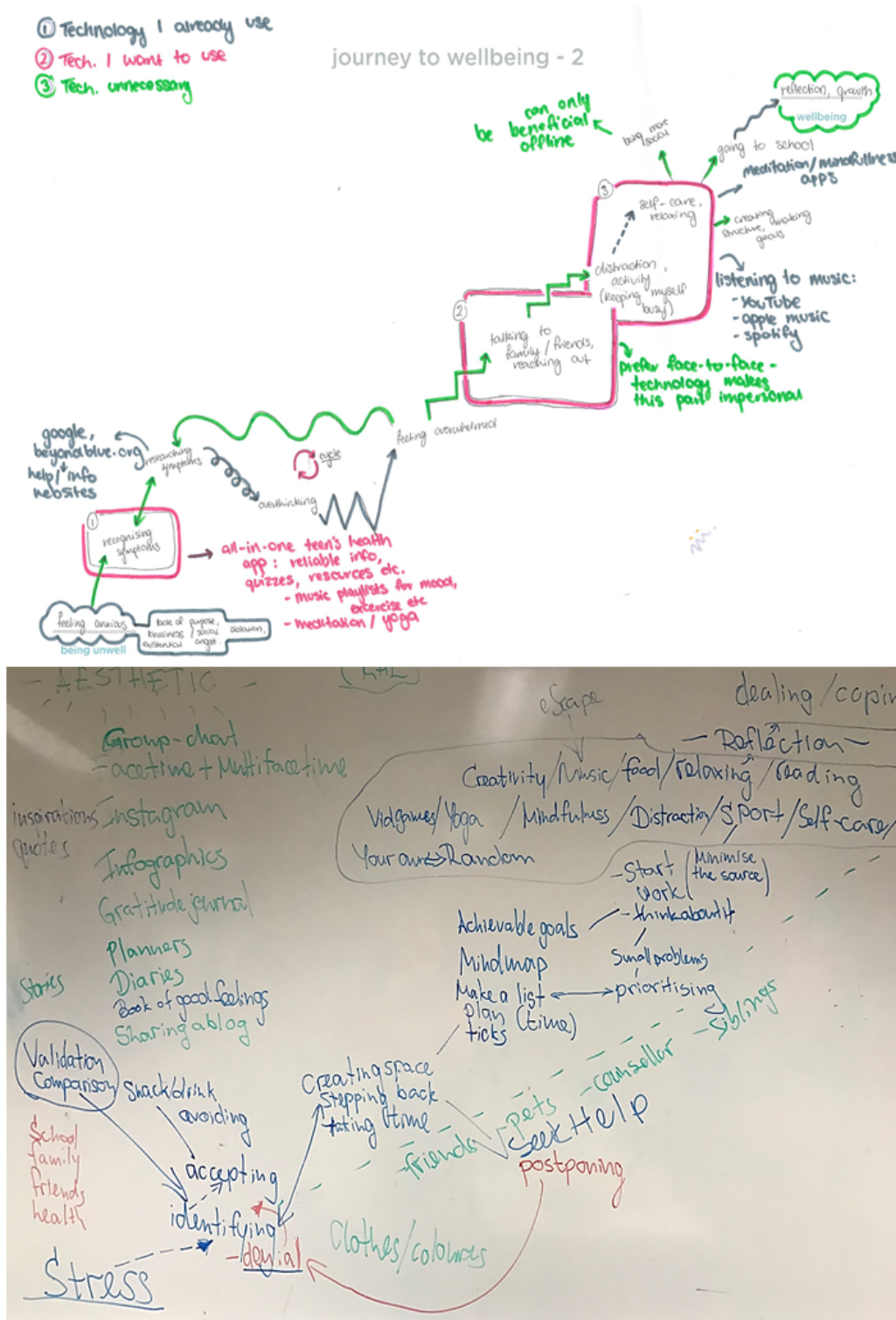
#### Step 3

Participants drew, on a blank sheet, a *recovery journey* depicting a hypothetical or experienced journey from being unwell to well and were provided with a specially designed selection of lines and arrows to mark the steps in meaningful ways (ie, indicate difficult steps, easy steps, barriers, etc). The team deemed it ethically appropriate to phrase journeys as hypothetical scenarios, to allow participants the flexibility to impersonally reveal potentially distressing, confronting, and stigmatizing personal experiences in front of the group. Participants engaged in continuous verbal annotations of their own maps, which was followed by a group synthesis of all journeys.

#### Step 4

Participants were asked to reflect on their journey maps and consider which steps could be assisted by technology and in what capacity (Figure 1). Thereafter, in-depth group discussions aimed to answer, "What type of technology would be best suited to provide assistance?" and "What technological functions would best address youth needs in progressing through the journey to well-being?" Participants were invited to work in pairs to create their most-desired mHealth tool (ie, with no limit to imagination and if resources were unlimited) to support the journey to well-being. Decisions to end data collection were based on saturation of emerging themes (Multimedia Appendix 2).

Figure 1. Individual and group journey maps.



Data Analysis

Analysis was largely informed by a social constructionist perspective [51], which assumes that people construct their own meaning of reality through interactions with others within a social space. This theoretical framework also allows for the interpretation of both linguistic and visual methods [49]. Hence, we viewed the data as forms of self-accounting in which participants were attentive to the expectations associated with

the production and reception of identities, both in immediate and broad social contexts.

Verbal data were transcribed verbatim and COREQ (Consolidated Criteria for Reporting Qualitative Research) [52] guidelines were followed. Codes were organized into thematic groups following inductive thematic analysis [53]. One team member with 7 years of experience conducting qualitative psychological research coded the transcripts in 3 iterative steps. After each coding step, a discussion was held with 2 additional

team members, who were also experienced in qualitative research, to reach agreement on the codes, themes, and theme names. Visual data were interpreted using photo elicitation analysis [42] and image clustering patterns were mapped over the verbal elaborations of the thematic analysis [53]. Comprehensive notes (on both verbal and visual data) were taken during the iterative process and incorporated in the final analysis.

### Intervention Design and Development

Aggregate technological solutions data were used to develop the concept design, storyboard, and functions of the mHealth tool, unanimously conceptualized by participants as a mobile app. The functions, information, and interface of the app were designed using iterative co-design—with young people and an expert development team. Principles outlined by the Mobile App Rating Scale (MARS) [35], which outline a set of quality criteria and consist of 29 items organized in 4 objective (engagement, functionality, aesthetics, and information quality) and 2 subjective (subjective quality and perceived impact) subscales, were followed. MARS provides a checklist of criteria and definitions to assist developers in creating high-quality health apps.

### The Role of the Researchers

The team consisted of researchers, designers, information technology developers, and young people. The collaborative approach served to ensure the thorough, interdisciplinary exploration of the psychological phenomena and the appropriate, grounded in theory, and participant data co-design, and development of the new technology.

## Results

### Participant Characteristics

A total of 25 participants residing in an urban setting in Australia took part in the study (Table 1). We planned to have each group be a mixed of genders (all participant identified as either male or female); however, 3 male participants who signed up for the 12-to-15-years-old age group did not attend on the day and did not respond to further communication attempts. Thus, this age group was only represented by female participants. Two 16-to-19-years-old groups were run in parallel because of a large amount of participant interest in that age range.



**Table 1.** Participant characteristics.

Characteristics	Age group		
	12 to 15 years (n=7), n	16 to 19 years (n=14), n	20 to 25 years (n=4), n
<b>Gender</b>			
Female	7	8	3
Male	0	6	1
<b>Ethnicity</b>			
White, Australian	6	6	3
White, New Zealander	0	1	0
Asian	0	5	1
Middle Eastern	1	1	0
European	0	1	0
<b>Education</b>			
Completed primary school	0	1	0
Partially completed high school	7	9	0
Completed year 12	0	4	2
Bachelor's degree	0	0	2
<b>Employment status</b>			
Full-time employment	0	0	0
Part-time or casual employment	0	4	1
Full-time student	7	4	2
Part-time student	0	1	0
Unemployed	0	5	1
<b>Living arrangement</b>			
With parents or guardians	7	12	2
With friends or in shared accommodations	0	2	1
Alone	0	0	1
<b>Relationship status</b>			
Single	7	13	3
In an exclusive relationship	0	1	1
<b>Currently receiving psychological treatment</b>			
No	6	13	2
Yes	1	1	1
Prefer not to say	0	0	1

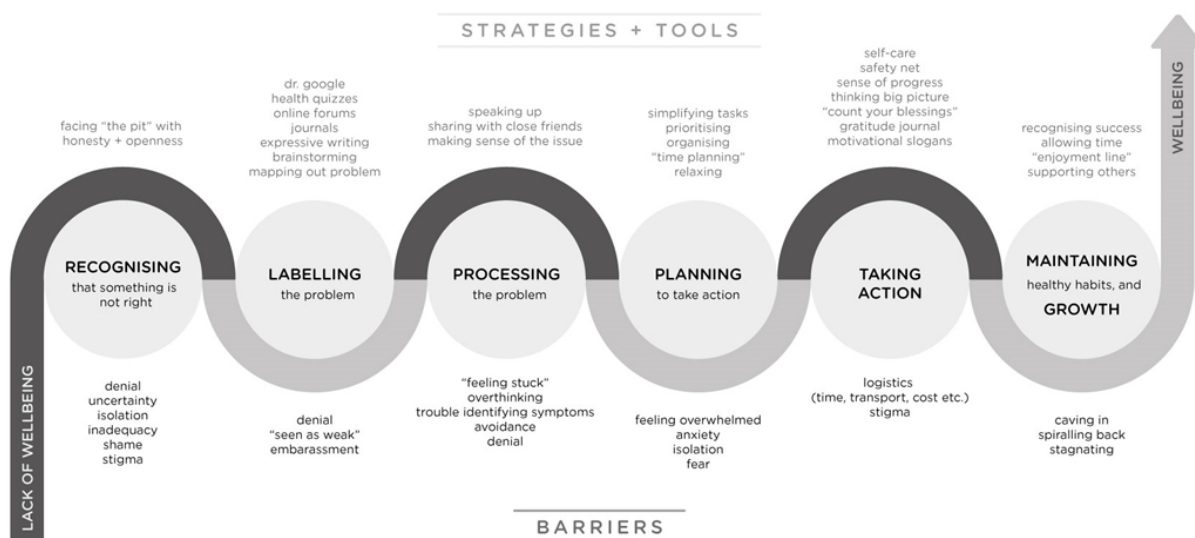
## Themes

### Overview

Young people indicated that the recovery journey was experienced as a complex, multifaceted phenomenon. Thematic analysis suggested that the journey consisted of 6 separate,

dynamically related stages (Figure 2). It commonly started as “depression,” “sadness,” “loneliness and isolation,” and “feeling fearful”, then progressed toward “feeling happy, loved, and accepted” or “overcoming fear and sadness.” The described journey stages significantly overlapped across all age groups and are presented here as forming a sequence, however, the journey could be nonlinear.

Figure 2. Young people’s journey toward well-being.



**Stage 1: Recognizing That Something Is Not Right**

This first stage was described as

*the pit...it's like a learning curve. So you get so deep into it and you don't really know how to climb back out. So it's like you're kind of like you want to recognise the problem but you don't want to recognise those steps cause you're scared or anxious.* [Participant, 16-to-19-years-old age group]

*You have to accept that something is wrong. I think a big part of it is you know something is wrong but being like "okay I need to get help for this. I can't do it on my own."* [Participant, 16-to-19-years-old age group]

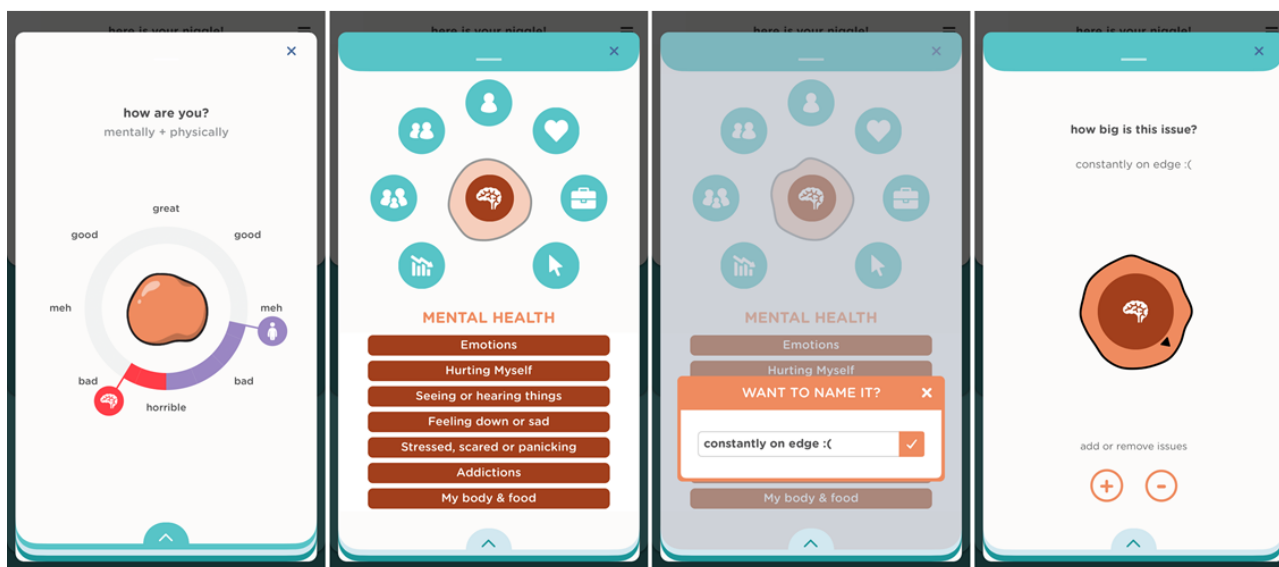
Barriers that were described included a sense of uncertainty, insecurity, and denial that lead to isolation and loneliness. Participants recognized that feelings of depression or anxiety

could result in experiencing shame, inadequacy, and stigma, and if left unacknowledged, such emotions may lead to further stagnation. Conversely, we gathered that facing the pit was a positive first step of accepting such feelings with honesty and openness to overcome "the negative spiral down."

*I feel like you just try and push it away and ignore it.... That'll stress me out even more.* [Participant, 12-to-15-years-old age group]

Participants suggested that the app should contain tools allowing reflection on the current physical and psychological state and a selection of common problems for their age to support this step of the recovery journey. The design solution included sliders used to modify the color of the amorphous problem, which the young people called "niggle," and a selection of issues from Kids Helpline’s youth problem classification lists (Figure 3).

Figure 3. Well-being log, labeling, and resizing the Niggle.



### Stage 2: Labeling the Problem

Young people could move onto accepting their niggle to get closer to overcoming it by recognizing it as a problem and giving it a name. Participants indicated that this stage was difficult due to the barriers of denial or avoidance.

*...you can still identify it and not want to do anything about it because you either don't have the energy to look into it, or you don't have the self-esteem to think you deserve better.* [Participant, 16-to-19-years-old age group]

Participants described useful strategies such as researching symptoms online, looking up health quizzes, or accessing online forums. They recognized that checking symptoms on Google was risky due to the lack of distinction between helpful and hazardous advice. Fleshing out the problem and giving it a name represented a greater sense of control, a broader perspective, better self-awareness, and acceptance. Participants discussed that acknowledging embarrassment, shyness, and nervousness at this stage was common and a key strategy used by participants to normalize their experience and continue the journey.

*That's pretty hard, like you would have to really sit down with yourself and say "alright, I don't feel comfortable doing this, but I think I might have to." You could look online and see ways you could like overcome pride... Yeah you have to put effort into identifying it means you want to do something about it which means you care about yourself.* [Participant, 12-to-15-years-old age group]

Successful tools, described by participants, for recognizing and labeling an issue included journaling, expressive writing, brainstorming, and mapping out the problem.

*Sometimes I just write down how I feel and like what's going through my mind and I go "what could this be?"* [Participant, 12-to-15-years-old age group]

To support this stage of the journey, the app depicts the problem as an amorphous niggle slowly taking shape. Users can choose up to 3 issues, give them their own labels, and adjust the size of each according to its gravity (Figure 3).

### Stage 3: Processing the Problem

The next stage of the journey toward well-being involved a contemplation period of stepping back and creating space in

order to process the confrontational experience of the realization that "something's not right" [Participant, 16-to-19-years-old age group]. This stage required persistence, deliberate reflection, and the ability to stay with the uncomfortable emotions. Young people spoke of prioritizing, rationalizing, and separating oneself from the problem in order to process it. They described recognition as being slow, "shaky," and nonlinear. A participant compared it to the "learning of a new complex skill, such as coding," which requires time and diligence and builds upon layers of newly acquired confidence.

*Wellbeing and having strong mental health is another skill you have to gradually build. Learning how to program websites, I've always had to refer back to all the different codes but now...I feel confident I have those codes...Once you get into those healthier habits...physically or intellectually you're more independent, your wellbeing will be better, and you will be able to move through the process.* [Participant, 20-to-25-years-old age group]

Barriers mentioned by the participants included feeling stuck, difficulties identifying the symptoms, overthinking, avoidance, or rushing for "quick fixes." To overcome these, participants relied on speaking up, educating themselves, sharing with close friends, and comparing to others with similar experiences, but not yet focusing on solutions.

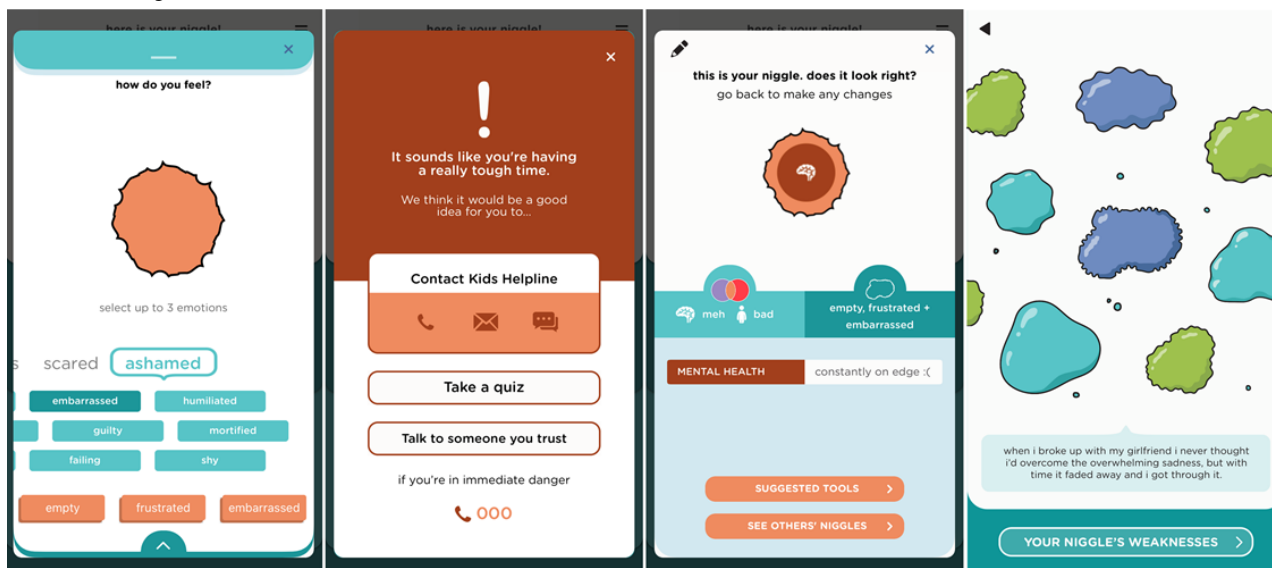
*Yeah, if it's a problem that you can't change, then stepping back is probably a good thing, but if it's a problem that you can change...* [Participant, 12-15-years-old age group]

*...stepping back I think has to do with like, identifying and then being able to think about priorities.* [Participant, 12-to-15-years-old age group]

*Yeah like um, separating yourself from an issue.* [Participant, 12-to-15-years-old age group]

To support progress through this stage, the app offers 2 elements. (1) The app invites users to take a moment to reflect on their emotions; they can select up to 3 emotions to associate with their current niggle, which changes its shape on the screen. (2) Once this 3-stage setup process is complete, it presents other users' niggles—a scrollable screen full of amorphous shapes with different colors and outlines. Tapping on any one of the shapes opens a support message from other users who have dealt with similar niggles (Figure 4).

Figure 4. Processing screens.



#### Stage 4: Planning Action

Participants shared that actively approaching a problem, learning more about it, brainstorming solutions and planning appropriate steps followed in the recovery journey once the processing stage was successfully completed and enough strength was gained. Planning involved setting achievable goals and a to-do list. Simplifying tasks by breaking them down into small steps, prioritizing and organizing a predictable timeline rather than dwelling on the big picture helped significantly. Rewards and “tick boxes” were regarded as positive reinforcement for sticking with the plan.

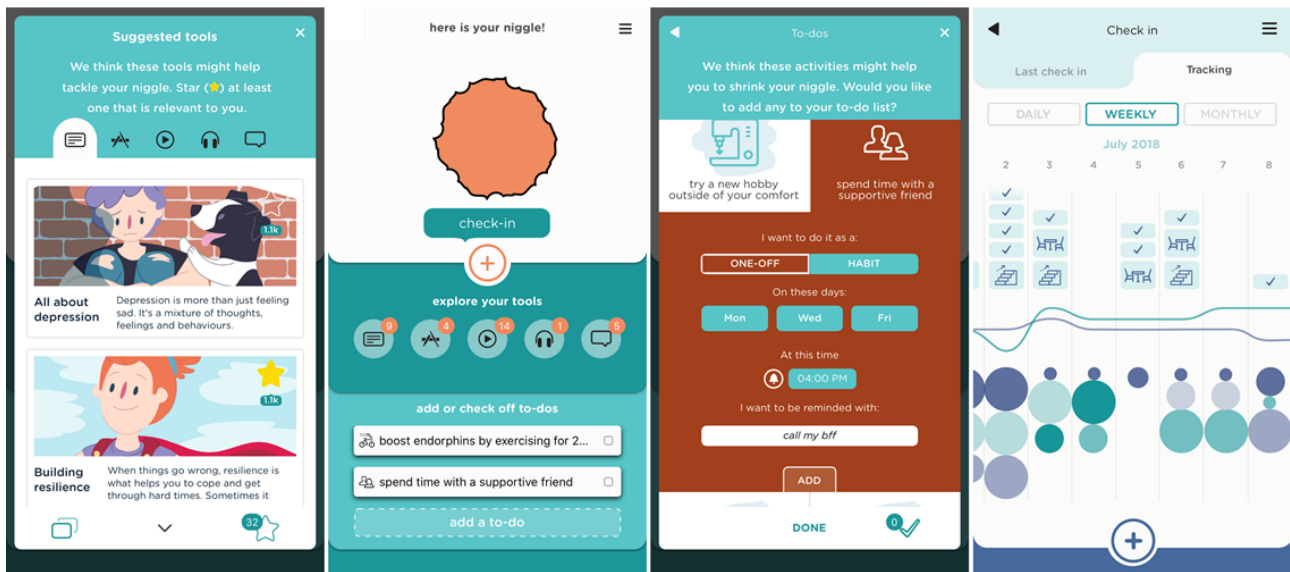
*When I'm overwhelmed I know that I like to make a list of my priorities and like, things that I need to do just to kind of like, see it and tick it off... [Participant, 12-to-15-years-old age group]*

*Yeah like kind of breaking down everything you need to do into smaller bits and like kind of, working through it slowly. [Participant 2, 12-to-15-years-old age group]*

Making time for “chilling or meditating,” listening to music, or engaging in pleasant hobbies were described by participants as additional tools in the arsenal of coping strategies. They mentioned barriers that included feeling overwhelmed by tasks, fear, increased anxiety, and emotional isolation. Some participants shared that planning could be a challenging and confronting experience, as it requires a greater realization of one’s concerns and the difficulty of overcoming them. Others mentioned their worries regarding help seeking such as distrust of the sources of support.

As learning and decision making characterized this step, the new app incorporates 2 key functions. (1) Users are presented with carefully researched or specially developed information about the selected issues in the form of tip sheets, videos, podcasts, relevant apps, and recovery stories from peers. Since searching the web for information can be problematic, resources were selected by a team of psychology students and reviewed by experts. (2) Young people are offered a list of helpful activities (to-do lists) specific to their niggles, which were developed by youth counselors. These activities can be scheduled directly into the app’s integrated calendar (Figure 5).

Figure 5. Learning and Planning screens.



### Stage 5: Taking Action

This stage was characterized as actively taking steps to resolve the problem. Young people felt particularly vulnerable at this stage of recovery. The intimidation of taking the first step, lack of accessibility to support, time limitations, and practical issues such as transportation and the cost of psychotherapy were all mentioned; however, participants indicated dealing with public and perceived stigma and shame (for speaking up about mental illness) was the most prevailing problem. Thus, escapism and quick fixes (such as unhealthy eating; bingeing on alcohol, drugs, shopping, or video games) were described as tempting alternatives.

*I think of escaping, it helps for a part. It helps the issues dull down and it helps you accept the issue and kind of deal with it. But after you escape you have to deal with the emotion.* [Participant, 20-to-25-years-old age group]

Helpful behaviors included speaking with family, friends, psychologists, counsellors, online or telephone support services or community groups (Figure 6). Technology was found to be both useful (online counseling, forums, playing video games, or using apps and websites) and unhelpful, particularly, the use of social media.

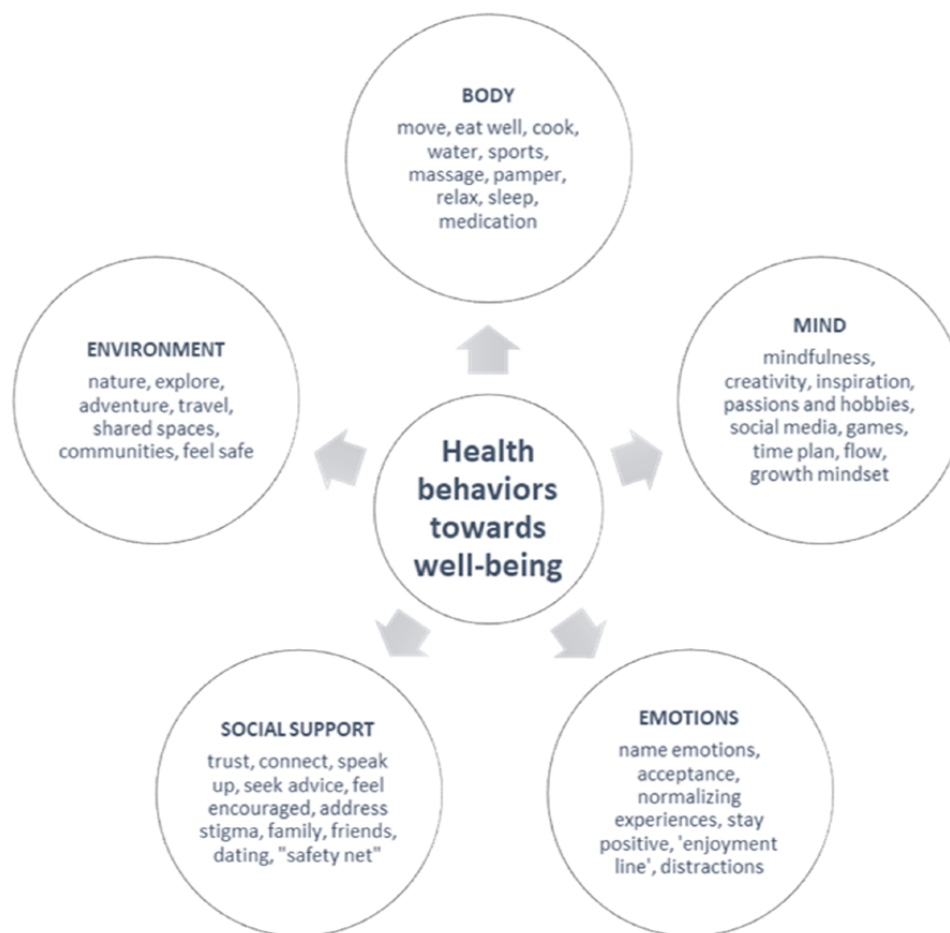
*...stay present and truly with friends rather than on phones...* [Participant, 16-to-19-years-old age group]

*Focus on stuff that you were able to do; and you feel proud of yourself... And from there, I put the gradual steps; to finding someone to talk about it... Because once you know that your feelings are valid, there's a part of your mind that wouldn't know what to do and that you might need help.... And then meanwhile, you can learn new things or find new hobbies that will occupy you.* [Participant, 20-to-25-years-old age group]

Positive relationships and trustworthy communities served as a powerful "safety net to catch you if you were to fall back into bad habits."

*It's not like it's a solo mission. You've got to love and trust the people that are there for you. Even if you feel like you're isolated or whatever, there are people that are there for you. So you have to put your trust in those people to feel like you've got the confidence to get help.* [Participant, 12-to-15-years-old age group]

At this stage of the journey, the app sends reminders for the scheduled to-do list tasks and allows users to add new activities, track their progress, and work toward creating habits. To minimize user effort, most activities can be modified and tracked with simple app interactions, such as tapping a checkbox to indicate a completed item, and are gamified using streaks and badges for increased engagement.

**Figure 6.** Health-enhancing behaviors for well-being.

### Stage 6: Maintaining Healthy Habits and Growth

Participants discussed that creating a positive structure around newly formed habits to maintain well-being constituted the final phase of the journey. Affirmation and recognition of the progress, usually within “the safety net” helped to prevent regression into earlier stages. Focusing on the enjoyment motivated young people to stay with the process at times of fallbacks. They described the “enjoyment line” as a

*“line between depression/happiness“; that you can have enjoyment; that enjoyment can actually bring you up. So if there's that line, when you reach a point where you actually have enjoyment again; so enjoyment can bring you further up; and then more care and love can bring you up to the point of wellbeing and happiness... [Participant, 16-to-19-years-old age group]*

*It's probably movable. It's not in a straight way. You're kind of like 'oh I've made so much progress on myself and I've come such a long way and I think that's good'. And it just boosts you up a bit more. [Participant, 16-to-19-years-old age group]*

This stage incorporated an additional aspect of growth. Participants' descriptions revealed that, having successfully taken the recovery journey, young people felt compelled to open up to others and offer advice, support, and help and that, at that

time, personal experience could be used to normalize and empower others' journeys to well-being.

This stage of the journey was supported in the app through (1) a sophisticated calendar tracking feature where users can monitor their progress and logs daily, weekly, or monthly; (2) gamification—achieving a new habit by completing a series of to-do items was rewarded by the *to-do* icon moving above the enjoyment line; and (3) inviting users to submit their personal recovery story of growth and their own niggles with other app users, thus closing the recovery cycle.

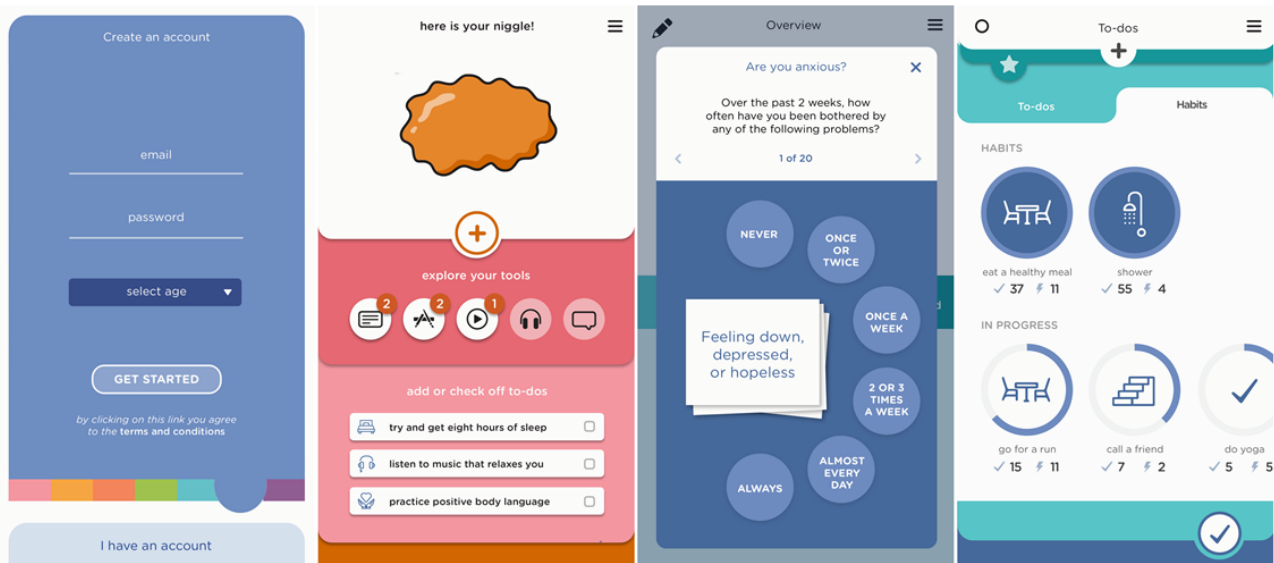
### The Niggle App

The app name was chosen by participants from the youngest age group (12-15 years) and was incorporated as a core concept of the app. Users are invited to reflect on their current physical and psychological well-being, identify their concerns or niggles, step back and reflect on their emotions, consider whether they need to discuss them with Kids Helpline, read about others' similar experiences, learn from a wealth of carefully selected information and resources, plan to-do lists, take action, and track their progress. Upon successfully completing the recovery journey, they can share their story with peers. There are additional app features to increase the customizability, usability, and engagement of the app. For example, users can select from 7 color templates or use quizzes to assess their distress, resilience, and well-being (Figure 7). A notable positive aspect of the recovery journey involves being able to access a network

of support from family, friends, and the community, which Niggle aimed to expand by creating an in-app community and including hotlink buttons to Kids Helpline as the first version of users' "safety net." Using the hotlink offers app users queue

priority for receiving immediate support. Multiple additional features were conceptualized based on the workshops and will gradually be implemented in future app updates.

**Figure 7.** App features to increase engagement.



## Discussion

### General

This paper presents the theoretical development and concept design phases of Niggle—a smartphone app to assist young people on every step of their recovery journey to well-being. App quality was ensured by following principles outlined by the MARS [35].

We grounded our research in the experiences of young people through participatory research [38,40,41,50] and co-design [37] to develop a clear understanding of their recovery journey and to allow young people to have more control over the process. While thematic analysis was rigorously applied to interpret data [53] and emerging themes were commonly supported by participants, we acknowledge that variability may exist regarding experiences in view of age, gender, ethnicity, race, and class differences. The young people who participated generally described the recovery journey as a multistage process, and their group discussions helped define 6 distinct stages. There was strong consensus that transition through these stages was nonlinear, and each stage was characterized by multiple barriers and helpful strategies. Therefore, Niggle corresponds to the intervention-delivery model through its nonlinear and highly customizable design.

Journey mapping allowed us to explore the barriers that young people encounter in help seeking and how they differ at each stage. An in-depth understanding of these concerns can assist health care providers in better addressing a young person's needs. For example, while denial, stigma, fearfulness, and avoidance seem to be present throughout the journey, specific logistics such as time, transportation, costs, and the availability of services are pertinent to the more active *planning action* and *taking action* stages.

Applying visual research methods allowed us to increase our confidence of overcoming the potential linguistic barriers that our youngest participants may have experienced. Thus, we present the useful strategies and tools young people rely upon to overcome barriers and how they can be translated into co-designed digital intervention features to support recovery.

Recovery has been theorized as being both an outcome and a process, and is still being debated in the literature [54]. Supporters of recovery as a process argue that this conceptualization as restoration of capacities to the stage preceding illness is partial and flawed, and our findings support this view. We found that recovery consists of the complementary experiences of restoration from mental illness and advancement through the stages of well-being (identifying, processing, and taking action, which results in growth). The themes that we identified largely align with those in descriptive psychological models of behavior change such as the transtheoretical model (TTM) [55]. The first stages of the TTM reflect a process of precontemplation (no intention to seek help in the near future) and contemplation (considering a change some time in the future based on the potential positives and negatives of that change) similar to those defined by our participants as *recognizing that something is not right* and *labeling the problem*). The subsequent TTM stages include preparation (an intention to seek help in the immediate future), corresponding to our *processing the problem* and *planning action* stages; the fourth TTM stage action (or currently seeking help) is similar to the *taking action* stage of the well-being journey. The TTM includes the stages of maintenance (help seeking when needed in an attempt to avoid relapse of problems) and termination (when help seeking is no longer required), which in our participants' discourse emerged as *maintaining health habits and growth*. Lastly, in line with our findings, the TTM also recognizes that relapse is a common

and normal part of the change cycle and defines 'recovery' as a process, rather than a state, or outcome.

Aligned with those of Prochaska et al [55], our findings, based on the participants' discussions, confirm that the recovery journey could be nonlinear. We argue that there is not a single, predetermined solution for dealing with each stage. While motivational work might assist with contemplation in the TTM, it may also be applicable to all stages of the journey. Similarly, skills training does not need to be limited to the action stage, as it may be appropriate for building confidence during contemplation or preparation. Our results reflect the TTM's applicability to young Australians in an era of unprecedented technological development, information overload, and shifting help-seeking pathways.

Consistent with previous findings [31], our findings also identified barriers such as stigma, embarrassment, poor mental health literacy, and a preference for self-reliance to youth help seeking, while facilitators included perceived positive past experiences, and social support.

### Limitations and Further Research

Despite efforts to capture diverse experiences, the participant sample was relatively homogenous. Most participants identified as female, urban residents, educated, White, living with parents or guardians, and single. None were employed full-time. We presume that young people from other backgrounds (such as Aboriginal and Torres Strait Islander, or culturally and linguistically diverse groups) would have provided a different account of recovery, mental health, self-care, and support safety nets. Yet, further work by our team, outside of the scope of the current project, is ongoing to confirm the relevancy of the stages reflected in the app design to a diverse group of young people, that includes young people from Indigenous backgrounds, young people with mental health diagnoses, and young people with diverse socioeconomic status.

The fact that male participants did not attend the workshop for the youngest age group (despite initially signing up), may reflect previous findings showing that the lowest rate of mental health help seeking is among young men, with only approximately 13% engaging with mental health services compared to 31% among young women [56,57]. This phenomenon has been linked to stigma (male role expectations and perceptions of appearing weak) [58], as well as low mental health literacy [31,56]. We hope that the Niggle app will be used to address some of these

issues by offering confidential access to support and psychoeducation to all users.

Using preselected photos to elicit discussions on well-being may have biased participant discussions; however, we believe that this photographic prompt served to open deeper discussions. Future research might benefit from using verbal and visual data in tandem (ie, both photo elicitation and photo voice [59]) and by encouraging young people to produce their own images.

App quality and efficacy evaluation projects, including feedback sessions with young people and a series of Randomized Controlled Trials will be carried out to determine the efficacy and effectiveness of Niggle.

### Implications for Practice

This study contributes to the field of well-being, mental illness, recovery, eHealth, and mHealth for young people. It demonstrates the usefulness of research into the alignment of theoretical models to specific demographics for the development of theoretically-sound digital health technologies. The recovery journey described by our participants confirms the applicability of the TTM [55] to support behavior change and well-being amongst young people in the context of eHealth. Despite international health policies increasingly placing well-being at their center [60], no agreement exists as to what constitutes well-being. The definitions presented in this paper, grounded in participants' complex personal accounts can assist with the development of policies and guidance for well-being and recovery interventions, regardless of the delivery method.

While there is a vast number of health apps targeting a range of users and issues, many lack theoretical grounding, quality, reliability, and applicability to key stakeholders. This paper offers an insight into health app development through in-depth phenomenological research, co-design, adherence to app quality principles, and the involvement of diverse expertise.

### Conclusion

We hope that Niggle will help young people access useful resources and support to increase their well-being. As of November 2019, version 1 of the app is freely available on Australian app stores. In the first year it was used by 12,823 individuals. Global release is planned once data privacy and security standards for international transfer and management have been carefully implemented.

---

### Acknowledgments

This project was primarily funded by the Australian Research Council (Linkage grant ARC LP150100178). Further financial contribution was provided by Kids Helpline and Queensland University of Technology. The authors of this paper would like to express their deepest gratitude to tall young people who took part in the project. Furthermore, we thank John Dagleish for his passion, support and contribution to this project and Tracy Adams for her dedication and commitment to the well-being of children and young people.

---

### Conflicts of Interest

None declared.

---



## Multimedia Appendix 1

Well-being workshops complete and numbered image set.

[[PDF File \(Adobe PDF File\), 3823 KB - mhealth\\_v9i4e21085\\_app1.pdf](#)]

## Multimedia Appendix 2

Topic guide.

[[DOCX File, 21 KB - mhealth\\_v9i4e21085\\_app2.docx](#)]

## References

1. Hides L, Dingle G, Quinn C, Stoyanov SR, Zelenko O, Tjondronegoro D, et al. Efficacy and outcomes of a music-based emotion regulation mobile app in distressed young people: randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Jan 16;7(1):e11482 [FREE Full text] [doi: [10.2196/11482](#)] [Medline: [30664457](#)]
2. Beauchaine TP. Future directions in emotion dysregulation and youth psychopathology. *J Clin Child Adolesc Psychol* 2015 May 27;44(5):875-896. [doi: [10.1080/15374416.2015.1038827](#)] [Medline: [26016727](#)]
3. Aldao A, Nolen-Hoeksema S, Schweizer S. Emotion-regulation strategies across psychopathology: a meta-analytic review. *Clin Psychol Rev* 2010 Mar;30(2):217-237. [doi: [10.1016/j.cpr.2009.11.004](#)] [Medline: [20015584](#)]
4. Laiho S. The psychological functions of music in adolescence. *Nord J Music Ther* 2009 Jul 10;13(1):47-63. [doi: [10.1080/08098130409478097](#)]
5. Dahl RE. Affect regulation, brain development, and behavioral/emotional health in adolescence. *CNS Spectr* 2001 Jan 07;6(1):60-72. [doi: [10.1017/s1092852900022884](#)] [Medline: [17008832](#)]
6. Dingle GA, Neves DDC, Alhadad SSJ, Hides L. Individual and interpersonal emotion regulation among adults with substance use disorders and matched controls. *Br J Clin Psychol* 2018 Jul 25;57(2):186-202. [doi: [10.1111/bjc.12168](#)] [Medline: [29277906](#)]
7. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jul 01;62(6):593-602. [doi: [10.1001/archpsyc.62.6.593](#)] [Medline: [15939837](#)]
8. Papinczak ZE, Dingle GA, Stoyanov SR, Hides L, Zelenko O. Young people's uses of music for well-being. *J Youth Stud* 2015 Mar 13;18(9):1119-1134. [doi: [10.1080/13676261.2015.1020935](#)]
9. Slade T, Johnston A, Oakley Browne MA, Andrews G, Whiteford H. 2007 National Survey of Mental Health and Wellbeing: methods and key findings. *Aust N Z J Psychiatry* 2009 Jul;43(7):594-605. [doi: [10.1080/00048670902970882](#)] [Medline: [19530016](#)]
10. Promoting mental health: concepts, emerging evidence, practice: a report of the World Health Organization, Department of Mental Health and Substance Abuse in collaboration with the Victorian Health Promotion Foundation and the University of Melbourne. World Health Organization. URL: [https://apps.who.int/iris/bitstream/handle/10665/43286/9241562943\\_eng.pdf?sequence=1&isAllowed=y](https://apps.who.int/iris/bitstream/handle/10665/43286/9241562943_eng.pdf?sequence=1&isAllowed=y) [accessed 2021-03-31]
11. Keyes CLM. Social well-being. *Soc Psychol Q* 1998 Jun;61(2):121. [doi: [10.2307/2787065](#)]
12. Hides L, Quinn C, Stoyanov SR, Cockshaw W, Mitchell T, Kavanagh DJ. Is the mental wellbeing of young Australians best represented by a single, multidimensional or bifactor model? *Psychiatry Res* 2016 Jul 30;241:1-7. [doi: [10.1016/j.psychres.2016.04.077](#)] [Medline: [27152903](#)]
13. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *Br J Psychiatry* 2011 Dec 02;199(6):445-452. [doi: [10.1192/bjp.bp.110.083733](#)] [Medline: [22130746](#)]
14. Keyes C. The subjective well-being of America's youth: toward a comprehensive assessment. *Adolesc Fam Health* 2006;4(1):3-11.
15. Keyes CLM. Mental health in adolescence: is America's youth flourishing? *Am J Orthopsychiatry* 2006 Jul;76(3):395-402. [doi: [10.1037/0002-9432.76.3.395](#)] [Medline: [16981819](#)]
16. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. *Lancet* 2015 Feb 14;385(9968):640-648 [FREE Full text] [doi: [10.1016/S0140-6736\(13\)61489-0](#)] [Medline: [25468152](#)]
17. Weiss LA, Westerhof GJ, Bohlmeijer ET. Can we increase psychological well-being? the effects of interventions on psychological well-being: a meta-analysis of randomized controlled trials. *PLoS One* 2016 Jun 21;11(6):e0158092 [FREE Full text] [doi: [10.1371/journal.pone.0158092](#)] [Medline: [27328124](#)]
18. Keyes CLM. Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. *Am Psychol* 2007;62(2):95-108. [doi: [10.1037/0003-066X.62.2.95](#)] [Medline: [17324035](#)]
19. Australian Psychological Society. 2015. URL: <https://www.psychology.org.au/getmedia/ae32e645-a4f0-4f7c-b3ce-dfd83237c281/stress-wellbeing-survey.pdf> [accessed 2020-03-17]
20. Redmond G, Skattebol J, Saunders P, Lietz P, Zizzo G, O'Grady E, et al. Are the kids alright? young Australians in their middle years. Australian Child Wellbeing Project.: Flinders University, University of New South Wales and Australian Council for Educational Research; 2016. URL: [www.australianchildwellbeing.com.au](http://www.australianchildwellbeing.com.au) [accessed 2020-03-17]

21. Eckersley R. A new narrative of young people's health and well-being. *J Youth Stud* 2011 Aug;14(5):627-638. [doi: [10.1080/13676261.2011.565043](https://doi.org/10.1080/13676261.2011.565043)]
22. Eckersley R. The health and well-being of young Australians: patterns, trends, explanations and responses. In: Bennett DL, Towns SJ, Elliott EL, Merrick J, editors. *Challenges in Adolescent Health: An Australian Perspective*. New York: Nova Science; 2009:3-19.
23. Lawrence D, Hafekost J, Johnson SE, Saw S, Buckingham WJ, Sawyer MG, et al. Key findings from the second Australian Child and Adolescent Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry* 2016 Sep 20;50(9):876-886. [doi: [10.1177/0004867415617836](https://doi.org/10.1177/0004867415617836)] [Medline: [26644606](https://pubmed.ncbi.nlm.nih.gov/26644606/)]
24. Mental health. Australian Institute of Health and Welfare. 2020 Jul 23. URL: <https://www.aihw.gov.au/reports/australias-health/mental-health> [accessed 2021-03-30]
25. Currie C, Zanotti C, Morgan A, Currie D, De Looze M, Roberts C, et al. Social determinants of health and well-being among young people. World Health Organization Regional Office for Europe. 2012 May. URL: [https://www.euro.who.int/\\_data/assets/pdf\\_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf](https://www.euro.who.int/_data/assets/pdf_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf) [accessed 2019-05-05]
26. Rickwood DJ, Telford NR, Parker AG, Tanti CJ, McGorry PD. headspace - Australia's innovation in youth mental health: who are the clients and why are they presenting? *Med J Aust* 2014 Feb 03;200(2):108-111. [doi: [10.5694/mja13.11235](https://doi.org/10.5694/mja13.11235)] [Medline: [24484115](https://pubmed.ncbi.nlm.nih.gov/24484115/)]
27. Price M, Hides L, Cockshaw W, Staneva A, Stoyanov S. Young love: romantic concerns and associated mental health issues among adolescent help-seekers. *Behav Sci (Basel)* 2016 May 06;6(2):9 [FREE Full text] [doi: [10.3390/bs6020009](https://doi.org/10.3390/bs6020009)] [Medline: [27164149](https://pubmed.ncbi.nlm.nih.gov/27164149/)]
28. Australian social trends. Australian Bureau of Statistics. 2007. URL: [https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/51EE403E951E7FDACA25732F001CAC21/\\$File/41020\\_Australian%20Social%20Trends\\_2007\\_ADJUSTED.pdf](https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/51EE403E951E7FDACA25732F001CAC21/$File/41020_Australian%20Social%20Trends_2007_ADJUSTED.pdf) [accessed 2019-05-05]
29. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? a systematic review of quantitative and qualitative studies. *Psychol Med* 2014 Feb 21;45(1):11-27. [doi: [10.1017/s0033291714000129](https://doi.org/10.1017/s0033291714000129)]
30. Yap MBH, Reavley N, Jorm AF. Where would young people seek help for mental disorders and what stops them? findings from an Australian national survey. *J Affect Disord* 2013 May;147(1-3):255-261. [doi: [10.1016/j.jad.2012.11.014](https://doi.org/10.1016/j.jad.2012.11.014)] [Medline: [23228570](https://pubmed.ncbi.nlm.nih.gov/23228570/)]
31. 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(1):1-9. [doi: [10.1186/1471-244x-10-113](https://doi.org/10.1186/1471-244x-10-113)]
32. Kids Helpline insights 2017: national statistical overview. yourtown. 2018. URL: <https://www.yourtown.com.au/sites/default/files/document/2017%20KHL%20Insights%20Report%20-%20Statistical%20Overview.pdf> [accessed 2019-05-05]
33. Kids Helpline insights 2018: national statistical overview. yourtown. 2019. URL: <https://www.yourtown.com.au/sites/default/files/document/KHL%20Insights%20National%20Statistical%20Overview%202018a.pdf> [accessed 2019-05-05]
34. Feroz A, Abrejo F, Ali SA, Nuruddin R, Saleem S. Using mobile phones to improve young people's sexual and reproductive health in low- and middle-income countries: a systematic review protocol to identify barriers, facilitators and reported interventions. *Syst Rev* 2019 May 18;8(1):117-117 [FREE Full text] [doi: [10.1186/s13643-019-1033-5](https://doi.org/10.1186/s13643-019-1033-5)] [Medline: [31103044](https://pubmed.ncbi.nlm.nih.gov/31103044/)]
35. 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](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
36. Mitchell KM, Holtz BE, McCarroll A. Patient-centered methods for designing and developing health information communication technologies: a systematic review. *Telemed J E Health* 2019 Nov 01;25(11):1012-1021. [doi: [10.1089/tmj.2018.0236](https://doi.org/10.1089/tmj.2018.0236)] [Medline: [30576258](https://pubmed.ncbi.nlm.nih.gov/30576258/)]
37. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, et al. Co-design of mhealth delivered interventions: a systematic review to assess key methods and processes. *Curr Nutr Rep* 2016 Jul 4;5(3):160-167. [doi: [10.1007/s13668-016-0165-7](https://doi.org/10.1007/s13668-016-0165-7)]
38. Hagen P, Collin P, Metcalf A, Nicholas M, Rahilly K, Swainston N. Participatory design of evidence-based online youth mental health promotion, intervention and treatment. Western Sydney University ResearchDirect.: Young and Well Cooperative Research Centre; 2012. URL: <https://researchdirect.westernsydney.edu.au/islandora/object/uws:18814/datastream/PDF/view> [accessed 2018-06-06]
39. Stoyanov S, Hides L, Kavanagh D, Sanders D, Cockshaw W, Mani M. A recommended process for development and evaluation of etools for mental health and wellbeing. In: *Innovations and Future Directions in the Behavioural and Cognitive Therapies*. Samford Valley, QLD: Australian Academic Press; 2016:251-255.
40. Robertson T, Simonsen J. Challenges and opportunities in contemporary participatory design. *Design Issues* 2012 Jul;28(3):3-9. [doi: [10.1162/desi\\_a\\_00157](https://doi.org/10.1162/desi_a_00157)]
41. Muller M. Participatory design: the third space in HCI. In: *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. NW Boca Raton, FL: CRC Press Inc; 2003:1051-1068.

42. Harper D. Talking about pictures: a case for photo elicitation. *Vis Stud* 2002 Jan;17(1):13-26. [doi: [10.1080/14725860220137345](https://doi.org/10.1080/14725860220137345)]
43. Richard VM, Lahman MK. Photo-elicitation: reflexivity on method, analysis, and graphic portraits. *Int J Res Method Educ* 2014 Apr 04;38(1):3-22. [doi: [10.1080/1743727x.2013.843073](https://doi.org/10.1080/1743727x.2013.843073)]
44. Howard T. Journey mapping. *Commun Des Q Rev* 2014 May;2(3):10-13. [doi: [10.1145/2644448.2644451](https://doi.org/10.1145/2644448.2644451)]
45. Tilleczeck KC, Bell BL, Munro MA, Gallant SD. Youth journeys in mental health: invoking unique patient perspectives. *Can J Commun Ment Health* 2017 Oct 01;36(2):151-164. [doi: [10.7870/cjcmh-2017-026](https://doi.org/10.7870/cjcmh-2017-026)]
46. Rose G. *Visual Methodologies: An Introduction to Researching With Visual Materials*. London: Sage Publications; 2012.
47. Emmison M, Smith P, Mayall M. *Researching the Visual*. London: Sage Publications; 2012.
48. Clark CD. The Autodriven interview: A photographic viewfinder into children's experience. *Visual Sociology* 1999 Jan;14(1):39-50. [doi: [10.1080/14725869908583801](https://doi.org/10.1080/14725869908583801)]
49. Croghan R, Griffin C, Hunter J, Phoenix A. Young people's constructions of self: notes on the use and analysis of the photo - elicitation methods. *Int J Res Method Educ* 2008 Oct;11(4):345-356. [doi: [10.1080/13645570701605707](https://doi.org/10.1080/13645570701605707)]
50. Carter B, Ford K. Researching children's health experiences: the place for participatory, child-centered, arts-based approaches. *Res Nurs Health* 2013 Feb 28;36(1):95-107. [doi: [10.1002/nur.21517](https://doi.org/10.1002/nur.21517)] [Medline: [23192941](https://pubmed.ncbi.nlm.nih.gov/23192941/)]
51. Burr V. *Social Constructionism*. East Essex: Routledge; 2015.
52. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec 16;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
53. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
54. Keyes C. *Mental Well-being: International Contributions to the Study of Positive Mental Health*. New York: Springer Science & Business Media; 2012.
55. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997;12(1):38-48. [doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38)] [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]
56. Rickwood D, Thomas. Conceptual measurement framework for help-seeking for mental health problems. *Psychol Res Behav Manag* 2012 Dec:173. [doi: [10.2147/prbm.s38707](https://doi.org/10.2147/prbm.s38707)]
57. Rodway C, Tham S, Ibrahim S, Turnbull P, Windfuhr K, Shaw J, et al. Suicide in children and young people in England: a consecutive case series. *The Lancet Psychiatry* 2016 Aug;3(8):751-759. [doi: [10.1016/s2215-0366\(16\)30094-3](https://doi.org/10.1016/s2215-0366(16)30094-3)]
58. Rice SM, Purcell R, McGorry PD. Adolescent and young adult male mental health: transforming system failures into proactive models of engagement. *J Adolesc Health* 2018 Mar;62(3S):S9-S17 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.07.024](https://doi.org/10.1016/j.jadohealth.2017.07.024)] [Medline: [29455724](https://pubmed.ncbi.nlm.nih.gov/29455724/)]
59. Wang C, Burris MA. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ Behav* 1997 Jun 01;24(3):369-387. [doi: [10.1177/109019819702400309](https://doi.org/10.1177/109019819702400309)] [Medline: [9158980](https://pubmed.ncbi.nlm.nih.gov/9158980/)]
60. Provencher HL, Keyes CL. Complete mental health recovery: bridging mental illness with positive mental health. *J Public Ment Health* 2011 Mar 15;10(1):57-69. [doi: [10.1108/17465721111134556](https://doi.org/10.1108/17465721111134556)]

## Abbreviations

**MARS:** Mobile App Rating Scale

**mHealth:** mobile health

**TTM:** transtheoretical model

*Edited by G Eysenbach; submitted 05.06.20; peer-reviewed by YH Kwan, S Sanatkar; comments to author 15.06.20; revised version received 11.09.20; accepted 31.01.21; published 20.04.21.*

*Please cite as:*

Stoyanov SR, Zelenko O, Staneva A, Kavanagh DJ, Smith C, Sade G, Cheers J, Hides L

*Development of the Niggle App for Supporting Young People on Their Dynamic Journey to Well-being: Co-design and Qualitative Research Study*

*JMIR Mhealth Uhealth* 2021;9(4):e21085

URL: <https://mhealth.jmir.org/2021/4/e21085>

doi: [10.2196/21085](https://doi.org/10.2196/21085)

PMID: [33877050](https://pubmed.ncbi.nlm.nih.gov/33877050/)

©Stoyan R Stoyanov, Oksana Zelenko, Aleksandra Staneva, David J Kavanagh, Calvin Smith, Gavin Sade, Jessica Cheers, Leanne Hides. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 20.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Results of a Culturally Tailored Smartphone-Delivered Physical Activity Intervention Among Midlife African American Women: Feasibility Trial

Rodney P Joseph<sup>1</sup>, PhD; Barbara E Ainsworth<sup>2,3</sup>, MPH, PhD; Kevin Hollingshead<sup>2</sup>, BSEET; Michael Todd<sup>4</sup>, PhD; Colleen Keller<sup>1</sup>, PhD

<sup>1</sup>Center for Health Promotion and Disease Prevention, Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

<sup>2</sup>College of Health Solutions, Arizona State University, Phoenix, AZ, United States

<sup>3</sup>Department of Kinesiology, Shanghai University of Sport, Shanghai Shi, China

<sup>4</sup>Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

**Corresponding Author:**

Rodney P Joseph, PhD

Center for Health Promotion and Disease Prevention

Edson College of Nursing and Health Innovation

Arizona State University

500 N 3rd St

Phoenix, AZ, 85004

United States

Phone: 1 602 496 0772

Fax: 1 602 496 1128

Email: [rodney.joseph@asu.edu](mailto:rodney.joseph@asu.edu)

## Abstract

**Background:** Regular aerobic physical activity (PA) is an important component of healthy aging. However, only 27%-40% of African American women achieve national PA guidelines. Available data also show a clear decline in PA as African American women transition from young adulthood (ie, 25-44 years) into midlife. This decline in PA during midlife coincides with an increased risk for African American women developing cardiometabolic disease conditions, including obesity, type 2 diabetes, and cardiovascular disease. Thus, effective efforts are needed to promote PA among sedentary African American women during midlife.

**Objective:** This study aims to examine the acceptability and feasibility of a culturally tailored, smartphone-delivered PA intervention, originally developed to increase PA among African American women aged 24-49 years, among a slightly older sample of midlife African American women aged 50-65 years.

**Methods:** A single-arm pretest-posttest study design was implemented. In total, 20 insufficiently active African American (ie,  $\leq 60$  min per week of PA) women between the ages of 50-65 years participated in the 4-month feasibility trial. The *Smart Walk* intervention was delivered through the study *Smart Walk* smartphone app and text messages. Features available on the *Smart Walk* app include personal profile pages, multimedia PA promotion modules, discussion board forums, and an activity tracking feature that integrates with Fitbit activity monitors. Self-reported PA and social cognitive theory mediators targeted by the intervention (ie, self-regulation, behavioral capability, outcome expectations, self-efficacy, and social support) were assessed at baseline and at 4 months. Feasibility and acceptability were assessed using a postintervention satisfaction survey that included multiple-choice and open-ended questions evaluating participant perceptions of the intervention and suggestions for intervention improvement. Wilcoxon signed-rank tests were used to examine pre- and postintervention changes in the PA and social cognitive theory variables. The effect size estimates were calculated using the Pearson *r* test statistic.

**Results:** Participants increased moderate-to-vigorous PA (median 30 minutes per week increase;  $r=0.503$ ;  $P=.002$ ) and reported improvements in 2 theoretical mediators (self-regulation:  $r=0.397$ ;  $P=.01$ ; behavioral capability:  $r=0.440$ ;  $P=.006$ ). Nearly all participants (14/15, 93% completing the satisfaction survey) indicated that they would recommend the intervention to a friend. Participants' suggestions for improving the intervention included enhancing the intervention's provisions of social support for PA.

**Conclusions:** The results provide preliminary support for the feasibility of the smartphone-based approach to increase PA among midlife African American women. However, before larger-scale implementation among midlife African American women, enhancements to the social support components of the intervention are warranted.

**Trial Registration:** ClinicalTrials.gov NCT04073355; <https://clinicaltrials.gov/ct2/show/NCT04073355>

(*JMIR Mhealth Uhealth* 2021;9(4):e27383) doi:[10.2196/27383](https://doi.org/10.2196/27383)

## KEYWORDS

exercise; physical activity; minority health; women's health; mHealth; mobile phone

## Introduction

### Background

African American women are disproportionately burdened by cardiometabolic diseases. Overall, 55% of African American women are obese [1], 57% have cardiovascular disease [2], and 13% have diabetes [3]. These cardiometabolic diseases affect African American women at higher rates than White and Hispanic women (ie, among non-Hispanic White women, 38% are obese, 43% have cardiovascular disease, and 7% have diabetes; prevalence rates among Hispanic women are 51%, 43%, and 12%, respectively). The high rates of these conditions among African American women contribute to their relatively high cardiovascular mortality rate, which is 2 times greater than that of non-Hispanic White and Hispanic women [4].

Recent data from 3 national data sets (ie, Behavioral Risk Factors Surveillance System, National Health and Nutrition Examination Survey, and National Health Interview Survey) indicate that only 27%-40% of African American women meet national aerobic physical activity (PA) guidelines (ie, 150 min per week of moderate-intensity PA, 75 min of vigorous PA, or an equivalent combination of durations and intensities [5,6]). These data also show a decline in PA as African American women transition from young adulthood (ie, 25-44 years) into midlife and older age [6]. The decline of PA among African American women during midlife is particularly concerning, given physiological changes associated with menopause (eg, decreased estrogen and progesterone production) that further increase a woman's risk of developing cardiometabolic diseases [7-9]. Thus, intervening to promote PA among sedentary African American women during midlife is critical to reducing cardiometabolic health disparities in this population.

Evidence accumulated over the past decade has shown that PA interventions delivered through mobile and smartphone devices (commonly referred to as mobile health [mHealth] PA interventions) are effective for promoting PA [10-12]. However, few mHealth PA interventions have focused on African American women [13-16], and even fewer have included midlife or older African American women [14,15]. This represents a missed public health opportunity, given the low PA levels, high cardiometabolic disease burden, and high level of smartphone use among African American women (recent data show that 79% of adults between the ages of 50 and 65 years own a smartphone, with limited or no differences by sex, race, or ethnicity [17]).

In a previous study, we developed *Smart Walk*, a 4-month culturally tailored, social cognitive theory (SCT)—based

smartphone-delivered intervention designed to increase PA and reduce cardiometabolic disease risk among African American women aged 24 to 49 years (ClinicalTrials.gov NCT02823379) [18]. *Smart Walk* was initially developed for women aged 24 to 49 years because it allowed us to tailor the PA intervention to the social, cultural, and behavioral characteristics of young to midlife African American women, who, at the time when the project was conceived (ie, in 2013), were the most likely age group to own a smartphone [19]. However, since *Smart Walk* was funded in 2015, smartphone use among midlife and older adults has increased substantially [17]. Similarly, recruitment efforts for the original *Smart Walk* pilot trial resulted in numerous women aged >50 years contacting our research team and expressing the desire to participate in the research study, indicating a demand for smartphone-delivered health promotion research among African American women aged ≥50 years.

### Objectives

This study aims to examine the feasibility and acceptability of the *Smart Walk* intervention among African American women aged 50-65 years. We hypothesized that data collected from the study would support the acceptability and feasibility of the approach and provide information on how the intervention can be refined to meet the social, cultural, and behavioral norms and preferences of women in this age group.

## Methods

### Study Design and Participants

This study was registered with ClinicalTrials.gov identifier NCT04073355. A single-arm pretest-posttest design was implemented. In total, 20 insufficiently active African American women with obesity (ie, BMI ≥30 kg/m<sup>2</sup>) aged between 50 and 65 years were recruited to participate in the 4-month smartphone-delivered PA intervention. Inclusion criteria were as follows: (1) self-identifying as African American and female, (2) aged 50-65 years, (3) having a BMI ≥30 kg/m<sup>2</sup>, (4) performing ≤60 minutes per week of moderate- to vigorous-intensity physical activity (MVPA) according to the 2-item Exercise Vital Sign Questionnaire [20], and (5) no self-reported ambulatory issues associated with moderate-intensity walking. Exclusion criteria included (1) concurrent participation in another PA, nutrition, or weight loss program at the time of enrollment or any time during the 4-month study and (2) indication of a potential contraindication of exercise according to the 2015 PA Readiness Questionnaire [21], unless a physician note allowing participation was provided. All study procedures were approved by the Institutional Review Board of Arizona State University.

## Description of the Smart Walk Intervention

*Smart Walk* is a culturally tailored, SCT—based 4-month PA intervention delivered through the *Smart Walk* smartphone app and text messages. The behavioral PA goal of the intervention was for participants to meet national guidelines of 150 minutes per week of at least moderate-intensity PA, with walking emphasized as the primary behavior to achieve this goal. An overview of the *Smart Walk* intervention is given below. Recent publications by our research team present in-depth descriptions of the development process [22] and design of the intervention [18].

### Smart Walk Smartphone App

The *Smart Walk* app includes 4 main features designed to promote daily PA: (1) personalized profile pages, (2) multimedia PA promotion modules delivered on a weekly basis in the form of brief videos and electronic text with images, (3) discussion boards for participants to discuss the weekly PA modules and give or receive social support, and (4) a PA self-monitoring or tracking tool that integrates with Fitbit (Fitbit Inc) activity monitors. Screenshots of app features were present in a study by Joseph et al [18].

### Personal Profile Pages

This feature was designed to be similar to profile pages available on commercial social media websites (eg, Facebook and Twitter) and provided participants with a platform to share select biographical information (ie, picture name, age, neighborhood or local area of residence, and brief biographical narrative) with other study participants. Profiles were designed to help facilitate building of a web-based community and social support network for PA.

### Multimedia PA Promotion Modules

Weekly multimedia video and text-based modules were the primary delivery channels for the educational and behavioral components of the program. These modules consisted of text- and image-based PA promotion materials written at less than eighth-grade reading level and brief 3- to 7-minute videos, narrated by the study's African American spokesperson (an African American woman in her mid-30s). Table 1 lists the module topics for the PA promotion program. During the first 3 months of the intervention, new PA promotion modules were delivered weekly. During month 4, new modules were presented every 2 weeks and focused on the maintenance of PA after the active intervention phase.

**Table 1.** Weekly physical activity topics covered during the intervention.

Week number	Module number	Module topic—PA <sup>a</sup> group
1	1	Introduction to the national PA guidelines and the health benefits of PA
2	2	Overview of PA-related health disparities among African American women and the importance of being a PA role model
3	3	Time management and strategies for incorporating 30 min of PA into the day
4	4	PA goal setting
5	5	Overcoming general barriers to PA
6	6	Tips for increasing daily PA
7	7	Overcoming hair care barriers to PA
8	8	Creating a social support network for PA
9	9	Trying new types of activities
10	10	Reducing sedentary time
11	11	Dietary behaviors to complement PA
12	12	Muscle strengthening and stretching activities to complement aerobic PA
13	N/A <sup>b</sup>	N/A
14	13	Dealing with setbacks
15	N/A	N/A
16	14	Review of previous modules and maintenance of PA after the active intervention phase

<sup>a</sup>PA: physical activity.

<sup>b</sup>N/A: not applicable.

### Discussion or Message Boards

Weekly PA promotion modules were accompanied by relevant discussion forum topics to encourage participants to reflect on the information presented in the multimedia PA promotion modules, share their personal experiences about PA, and give and receive social support for PA. The discussion board feature

also included a general *Community Board* forum and a *Meet-up* forum, where participants could share information and/or discuss topics that may not clearly align with the weekly module topics and discuss or coordinate group-based exercise activities. The dialog on these discussions was a primary mechanism through which social support for PA was fostered among participants. Push notifications were sent to study participants when other

participants posted on the discussion boards. Example module discussion board topics are as follows: “Why do you want to be physically active?” and “What are ways you add more physical activity into your day?”

### ***PA Self-Monitoring or Goal Setting Feature***

Participants received a wrist-worn Fitbit Inspire HR (Fitbit Inc) activity monitor to wear throughout the study. Data collected from the Fitbit integrates with the *Smart Walk* app to allow participants to view and track the minutes of MVPA performed during the study. The criterion used to define MVPA, as measured by the Fitbit, was walking at a cadence of 100 steps per minute [23,24] for at least 1 minute. Activity not meeting this intensity-duration criterion was not registered as MVPA on the *Smart Walk* app. For activities not recorded by the activity monitor (eg, stationary cycling, water aerobics, and swimming), participants could manually enter the activity via the app’s tracking feature. Fitbits were registered to a research study account to allow for staff to monitor PA progress and troubleshoot any device-related issues experienced by participants.

### ***PA Promotion Text Messages***

Participants received 3 PA promotional text messages each week. These text messages provided participants with inspirational quotes, reminders, and tips for increasing daily PA. Message content was developed through our formative research with African American women [25,26]. Example text messages are as follows: “Commit to being fit. It’s never too late to achieve your physical activity goals.” and “Be an active role model to those around you!”

### ***Theoretical Basis of the Intervention***

SCT [27] served as the theoretical framework for the intervention. Developed by Bandura [27], SCT posits that human behavior is a result of the reciprocal and dynamic interaction of personal factors and the social environment. Intervention components were designed to engage 5 constructs of SCT: behavioral capability, social support, self-efficacy, outcome expectations, and self-regulation. [Textbox 1](#) provides a brief description of each of these constructs and how the intervention was designed to leverage each construct to promote PA.



**Textbox 1.** Social cognitive theory constructs targeted by the intervention.

Behavioral capability—knowledge and skill to perform a physical activity (PA)

- Multimedia modules provide information on:
  - Definitions of PA and exercise
  - Different types of PA (ie, aerobic vs muscle strengthening)
  - National PA guidelines and the types of PA that can be performed to achieve the PA guideline
  - How to determine the intensity of PA performed

Social support—extent to which significant referents approve, encourage, and/or influence performance of PA

- Discussion board prompts facilitate and encourage participants to give and receive emotional support for PA
- Multimedia modules and text messages:
  - Provide participants with encouragement and empowerment for PA
  - Emphasize PA is a form of self-care and that African American women are worthy of self-care activities
  - Encourage participants to reflect on why they should be physically active
  - Include talking points and negotiation strategies to facilitate social support for PA

Self-efficacy—confidence in oneself to take action and overcome barriers. Self-efficacy is derived through 4 main sources: (1) Mastery experiences (ie, first-hand experience with performing a behavior), (2) Social modeling (ie, seeing other similar to ourselves successfully perform a behavior), (3) Emotional arousal (ie, improving emotional states by reducing stress or anxiety and by promoting positive emotions), and (4) Verbal persuasion (ie, verbal encouragement to engage in a behavior):

- Participants track PA using the activity tracking feature, allowing participants to enact PA strategies encouraged by the intervention and track their increases or decreases in PA (ie, mastery experiences)
- Multimedia modules:
  - Illustrate African American women engaging in various types of aerobic PA and include testimonials from African American women describing how and why they are physically active (ie, social modeling)
  - Reinforce the idea that PA does not have to be structured or difficult by encouraging walking and providing tips on how more walking can be incorporated into day (ie, emotional arousal)
- Multimedia modules and text messages encourage participants to be active by providing words of encouragement and empowerment for PA (ie, verbal persuasion)
- Discussion boards and weekly prompts provide a venue for participants encourage each other to be physically active (ie, verbal persuasion)

Outcome expectations—anticipated outcomes of engaging in PA

- Weekly video and text modules provide participants with the health and social outcomes associated with being physically activity, including:
  - Reduced risk for heart disease and type 2 diabetes
  - Weight maintenance
  - More energy to perform daily activities
  - Improved quality and quantity (ie, years) of life
  - Being a good role model to others

Self-regulation—ability to manage social, cognitive, and motivational processes to achieve a desired goal

- Participants use Fitbit activity monitor and *Smart Walk* PA self-monitoring feature to track daily and weekly PA.
- Intervention materials encouraged to achieve the static goal of 150 min per week of moderate- to vigorous-intensity PA. In addition, the week 4 module encourages participants to create short- and long-term goals associated with PA performance.
- The week 4 module encourages participants to create self-rewards for achieving previously established PA goals.

**Cultural Tailoring of the Intervention**

Cultural tailoring of the intervention was informed through an extensive review of the literature and our formative focus group

research with adult African American women (mean age of focus group participants was 38.5 years, SD 7.8). The intervention was designed to be sensitive to, and to leverage, the lived experiences, sociocultural norms, and familial and

societal expectations of African American women for the successful promotion of PA. On the basis of the cultural tailoring framework by Resnicow et al [28], the intervention addressed both *surface* and *deep-structure* cultural characteristics of African American women. Surface-level cultural tailoring refers to the most basic level of cultural tailoring and includes matching the characteristics and packaging of a promotion to the overt social and behavioral characteristics of the intended population [28]. Surface-level cultural tailoring of the *Smart Walk* was achieved through (1) evidential statistics emphasizing the low levels of PA and high prevalence of cardiometabolic disease conditions among African American women (eg, “Only 36% of African American women meet the National PA Guidelines” and “Almost half (49%) of African American

women over the age of 20 have heart disease”), (2) images of African American women with diverse physical characteristics (body shapes, hairstyles, and skin tones) performing PA throughout PA module text, and (3) having a local African American woman serve as the study spokesperson in the video modules. Deep-structure cultural tailoring involves recognizing sociocultural norms and values, beliefs, and behaviors of a group and using these characteristics to motivate behavior change [28]. Deep-structure cultural tailoring efforts focused on 3 key concepts: (1) collectivism or ethics of care, (2) physical appearance norms (ie, hair care or body shape concerns), and (3) racial pride or role modeling. Table 2 describes each cultural consideration and how the intervention was designed to influence these characteristics to promote PA.

**Table 2.** Deep-structure cultural tailoring characteristics of the intervention.

Cultural consideration	Brief description	How the cultural characteristic was addressed in the intervention
Collectivism	Prioritizing the needs and well-being of others (ie, family or friends) over the needs and well-being of oneself. This can contribute to African American women reporting lack of time, energy, or resources for PA <sup>a</sup> . Although this perspective has been reported among women of other races or ethnicities, previous research, including our own pilot work, suggests this phenomenon may be more accentuated in the African American community.	Intervention materials: <ul style="list-style-type: none"> <li>recognized the importance of caretaking in the value system of African American women.</li> <li>emphasized PA is an investment in the health and well-being of African American women and not a competing interest with caretaking, familial, or other responsibilities.</li> <li>portrayed regular PA as a key behavior to help participants perform their caretaking, familial, and community responsibilities with more energy and for a longer duration throughout the life span.</li> </ul>
Racial pride or role modeling	Awareness and interest in how one’s behavior can contribute the collective health and well-being of the African American community.	Intervention materials highlighted that physically active African American women are positive role models to other members of the African American community, which can encourage others in their community (ie, family and friends) to adopt a physically active lifestyle.
Physical appearance preferences	Some African American women are hesitant to engage in PA because <ul style="list-style-type: none"> <li>perspiration or sweat negatively impacts their hairstyle.</li> <li>they perceive PA will alter their desired body shape.</li> </ul>	Intervention materials: <ul style="list-style-type: none"> <li>included hairstyling strategies to help reduce the negative effects of perspiration (ie, use of hair wraps and dry shampoo strategies).</li> <li>encouraged women to adopt hairstyles that are less impacted by perspiration (ie, braids and natural hairstyles).</li> <li>informed participants that engaging in PA at the levels recommended by the study (ie, 150 min per week) will not substantially change their body shape unless they also change their dietary habits.</li> <li>emphasized health benefits of PA independent of weight loss (ie, reduced cardiometabolic disease risk, weight maintenance, and increased energy).</li> </ul>

<sup>a</sup>PA: physical activity.

## Measures

### Demographic Characteristics

Demographic characteristics (age, education, income, ethnicity, number of children in household, and marital or relationship status) were assessed using a self-administered questionnaire.

### Anthropometrics

Weight, height, and waist circumference were measured by a member of the research team in a private room while wearing light clothing but no shoes. Body weight was measured in kilograms using the Tanita TBF-300A (Tanita Corporation) digital scale. Height was measured in centimeters using a Seca 213 portable stadiometer. BMI was calculated as weight in kilograms divided by height squared (in meters).

### PA Outcomes

Self-reported weekly minutes of MVPA was assessed using the 2-item Exercise Vital Sign Questionnaire [20], and weekly energy expenditure was assessed using the REGICOR questionnaire [29]. The Exercise Vital Sign Questionnaire assesses the frequency (days per week) and duration (min per day) of MVPA performed (eg, a brisk walk) during the past week. The questionnaire is scored by multiplying the days by minutes per day of PA to generate an estimate of minutes per week of MVPA. This measure has been validated against population-based surveillance surveys [20] and accelerometers for the accurate assessment of PA [30]. The REGICOR questionnaire is a 16-item questionnaire that assesses leisure-time PA (including time spent in active transportation to work, walking, climbing stairs, and playing sports), time

sedentary behaviors, and a ranking of occupational PA. This provides an estimation of total energy expenditure as well as energy expended in light-intensity (<4 metabolic equivalent [MET]), moderate-intensity (4-5.5 MET), and vigorous-intensity (>5.5 MET) activities. Energy expenditure for leisure-time PA is expressed as MET-minute obtained by multiplying the MET intensity level by minutes reported for the activities. The scoring algorithm was described by Molina et al [29]. Here, we only report the results for leisure-time PA, as the study was not designed to target sedentary behavior or occupational PA. The REGICOR has high test-retest reliability for assessing MVPA (interclass correlation=0.79 and 0.95 for moderate- and vigorous-intensity PA, respectively), is sensitive to assessing changes in PA over time, and has been validated against accelerometers [29].

### **SCT Outcomes**

SCT constructs were assessed using a series of questionnaires developed and tested in previous research. Self-efficacy for PA was assessed using the 12-item Exercise Confidence Survey [31]. Social support for exercise was evaluated using the 20-item Social Support for Exercise Survey [32]. This scale consists of separate 10-item measures of family support and friend support. Self-regulation for PA was measured using the 10-item Self-Regulation Scale from the Health Beliefs Survey [33,34]. Outcome expectations for PA were assessed using the 9-item Outcome Expectation Scale for Exercise [35]. Behavioral capability for PA was measured using a 5-item scale developed by our research team, adapted from previous research [36,37] to assess knowledge of national PA guidelines and the health benefits of PA.

### **Acceptability and Feasibility**

Acceptability and feasibility were assessed by examining objectively measured indices of Fitbit wear, app use, and text message delivery and by evaluating self-reported treatment acceptance via a consumer satisfaction survey. Fitbit wear was assessed by examining the daily Fitbit wear. The criterion used to indicate daily wear was modeled after work by Hartman et al [38] and required participants to accumulate at least 1 minute of light- or moderate-intensity PA (as classified via Fitbit's tracking algorithm) on a given day. This criterion was selected because participants were instructed to wear the activity monitor to track their activity, as opposed to wear the device all day. Accordingly, achieving 1 minute of activity would indicate that the activity monitor was worn for at least part of the day. App use was assessed using analytic tracking software built into the app during development. Specific data points (events) recorded by this software included (1) when a participant loaded a PA module on her phone, (2) when a participant played a module, and (3) when a participant posted in a discussion board forum. Weekly text messages were sent to participants using Twilio, a commercial cloud-based communication platform, which integrated with the smartphone app platform. The app platform recorded the time and date when a text message was delivered to a participant. In the event that a text message was not delivered, an error note was sent to the research team. Treatment acceptance was assessed using a 37-item consumer satisfaction survey adapted from previous research [13,39]. This survey

included both multiple-choice and open-ended questions evaluating participants' perceptions of the intervention, including content and app usability, and suggestions for how the research team could improve the intervention. Example questions include "Overall, how helpful did you find the weekly videos and text modules to promote physical activity?" and "If we were to do the program again, what do you recommend we change, add, or remove from the program?"

### **Procedures**

Community-based recruitment strategies were used to recruit participants, including social media posts, newspaper advertisements, in-person recruitment at community events, and word-of-mouth. Women interested in study participation contacted the research team via email, telephone, or the study website (which was included in recruitment materials) and completed an eligibility screening survey by telephone or via the internet. Eligible women were scheduled to attend an in-person study orientation session held on the downtown campus of Arizona State University. This orientation session was designed to provide detailed information about the study activities and allow participants to have any questions regarding the study answered by the study staff. At the conclusion of the orientation session, women who indicated a desire to participate provided written informed consent; completed paper-based baseline surveys; and had their resting blood pressure, height, and weight assessed by study staff. To reduce participant wait time between baseline assessments and initiation of the intervention, participants were enrolled in 3 separate cohorts: cohort 1 included 9 participants who received the intervention from October 2019 to February 2020, cohort 2 included 5 participants enrolled from November 2019 to March 2020, and cohort 3 included 6 participants who received the intervention from December 2019 to April 2020. Before the start of each cohort, a member of the study team provided participants with a Fitbit activity monitor to wear on their nondominant wrist during the intervention and assisted each participant with downloading the *Smart Walk* app. Study staff proactively followed up participants every 2 weeks during the intervention to identify any issues participants may have with the *Smart Walk* app or Fitbit activity monitor. This contact also served as an opportunity to remind participants to wear the Fitbit daily.

Follow-up study assessment procedures were designed to be similar to baseline. Specifically, participants were expected to attend in-person assessments to complete the study surveys and have their weight and resting blood pressure assessed. However, in-person assessments were not possible for cohorts 2 and 3 because of safety concerns associated with the transmission of SARS-CoV-2 (COVID-19). Accordingly, participants in these 2 cohorts were mailed the 4-month survey packet. After completing the surveys, participants were asked to return the study materials to the research team using a prepaid mailer. However, several participants reported that they were unable to mail the study packet to the research team because they either lived in residences that did not have a mailbox large enough to place the prepaid mailing envelope and/or were unable to access a postal service dropbox. For these individuals, a member of the research team coordinated a no-contact pick-up of the survey and accelerometer at the participant's residence. Due to the

remote nature of follow-up data collection procedures for cohorts 2 and 3, blood pressure and anthropometric data were not collected for these participants. Accordingly, these data are not presented at follow-up because of the large amount of missing data at the 4-month assessment. Participants were provided US \$50 for study participation (US \$25 after the baseline assessment and US \$25 after the 4-month assessment) and were allowed to keep the Fitbit provided to them during the intervention.

### Sample Size Considerations

This study focused on examining the feasibility of the smartphone-delivered intervention among a sample of midlife African American women, rather than efficacy testing. We selected a sample size of 20 because it would yield sufficient information regarding the feasibility of implementing the intervention with a sample of midlife African American women and allow for a preliminary examination of pre- versus postintervention differences in outcomes.

### Statistical Analysis

Descriptive statistics (mean, median, percentage, and frequency) were used to summarize demographic characteristics and postintervention consumer satisfaction survey responses. Due to the small sample size and concerns about data not meeting distributional assumptions for parametric statistical tests, nonparametric techniques were used to examine baseline to 4-month changes in SCT and PA outcomes. With baseline values carried forward for missing data at follow-up, Wilcoxon signed-rank tests were used to examine pre- versus postintervention changes in study outcomes, and correlation coefficients were used to characterize the magnitude of change. Statistical significance was set at a  $P$  value of  $<.05$ ; however,  $P$  values are provided for reference only, as the study was not powered to detect significant changes in study outcomes. Qualitative data provided by participants on the consumer

satisfaction survey were analyzed using direct content analysis [40]. For this analysis, participants' responses to open-ended survey questions were initially coded based on the specific intervention component of the qualitative narrative referenced (multimedia modules, discussion boards, Fitbit, activity tracker, and text messages). Narratives not focused on a specific intervention component were coded into a category broad category entitled *overall satisfaction with the intervention*.

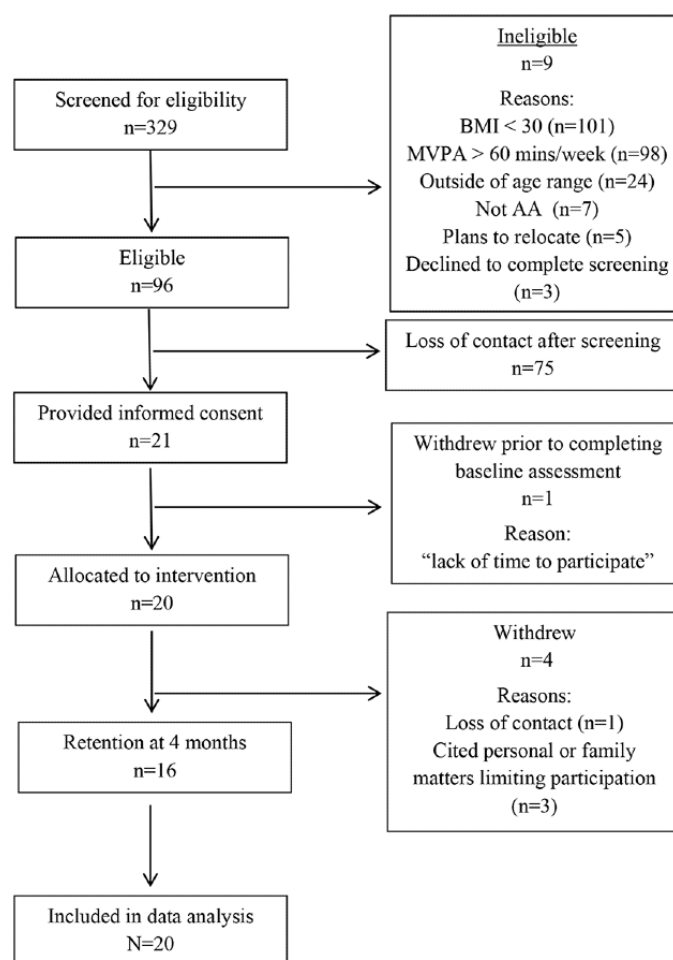
Next, narratives within each code were reviewed, and repetitive themes emerging from the coded data were used to reflect participant sentiments regarding the overall intervention as well as specific intervention components. A qualitative analysis was performed by one member of the research team (RPJ).

## Results

### Participant Flow and Baseline Characteristics

Recruitment efforts resulted in 329 women being screened for eligibility. Of these, 96 women were eligible, and 21 provided informed consent to participate in the study. The low recruitment rate (22%) from eligibility to enrollment was due to individuals completing the web-based eligibility screening survey advertised on social media but not responding to follow-up inquiries from the research team for enrollment. One participant withdrew from the study before receiving the intervention, citing a lack of time to participate, resulting in 20 women receiving the intervention. In total, 16 participants provided follow-up data at 4 months (ie, 80% retention), with 1 participant's completed follow-up data forms being lost by the project staff. The reasons for withdrawal at follow-up included loss of contact ( $n=2$ ) and personal issues limiting study participation ( $n=2$ ). [Figure 1](#) provides detailed information regarding participant flow throughout the study, including reasons for ineligibility and withdrawal from the study.

**Figure 1.** CONSORT (Consolidated Standards of Reporting Trials) diagram illustrating participant flow. AA: African American; MVPA: moderate-to vigorous-intensity physical activity.



Participants who began the intervention ( $n=20$ ) had a mean age of 56.2 (SD 4.3) years and a mean BMI of 40.0 (SD 8.6)  $\text{kg}/\text{m}^2$ . Overall, 45% (9/20) of women were married ( $n=8$ ) or living in a marriage-like relationship ( $n=1$ ), 30% (6/20) were either divorced ( $n=5$ ) or separated ( $n=1$ ), and 25% (5/20) were never married ( $n=5$ ). In addition, 40% (8/20) of women attended some college or technical school, 35% (7/20) had a bachelor's degree, and 25% (5/20) had a master's degree. One participant reported an annual household income of <US \$25,000, 30% (6/20) reported annual household incomes of US \$25,000-US \$50,000, 45% (9/20) reported incomes of US \$50,001-US \$100,000, and 20% (4/20) reported incomes >US \$100,000.

### PA Outcomes

The PA outcomes are presented in Table 3. On the basis of Exercise Vital Sign Questionnaire, self-reported minutes per week of MVPA increased from a median of 20 minutes per week at baseline to 50 minutes per week at 4 months ( $r=0.503$ ;  $P<.001$ ). Weekly estimated energy expenditure assessed by the REGICOR indicated comparable increases for total weekly energy expenditure (baseline median=58.04 MET-min per week; 4-month median=265.91 MET-min per week;  $r=0.407$ ;  $P=.008$ ), with changes in light-intensity ( $r=0.423$ ;  $P=.005$ ) and moderate-intensity activities ( $r=0.512$ ;  $P<.001$ ) accounting the majority of this change.

**Table 3.** Self-reported physical activity outcomes.

Variable	Baseline		4 months		P value <sup>a</sup>	Effect size <sup>b</sup>
	Mean (SD)	Median (range)	Mean (SD)	Median (range)		
Exercise vital sign (min per week)	27.00 (26.57)	20 (0-90)	65.50 (61.43)	50.00 (0-200)	<.001	0.503 <sup>c</sup>
REGICOR light energy expenditure (MET <sup>d</sup> -min per week)	36.83 (27.64)	37.30 (0-111.89)	84.15 (71.86)	55.94 (0-268.07)	.005	0.423 <sup>e</sup>
REGICOR moderate energy expenditure (MET-min per week)	27.39 (60.18)	0 (0-268.07)	86.66 (84.17)	67.02 (0-267.07)	<.001	0.512 <sup>c</sup>
REGICOR vigorous energy expenditure (MET-min per week)	78.05 (140.91)	6.12 (0-512.82)	93.46 (135.64)	9.03 (0-480.42)	.50	0.114 <sup>f</sup>
REGICOR total energy expenditure (MET-min per week)	142.27 (178.20)	58.04 (0-585.08)	264.26 (199.62)	265.91 (41.96-732.17)	.008	0.407 <sup>e</sup>

<sup>a</sup>Wilcoxon signed-rank *P* value for baseline to 4-month change.

<sup>b</sup>Pearson *r* effect size estimate.

<sup>c</sup>Effect size: large (0.50-1).

<sup>d</sup>MET: metabolic equivalent.

<sup>e</sup>Effect size: medium (0.30-0.49).

<sup>f</sup>Effect size: small (0.10-0.29).

### SCT Outcomes

Pre-post intervention changes in SCT constructs targeted by the intervention are presented in [Table 4](#). The results showed enhancements in self-regulation ( $r=0.397$ ;  $P=.01$ ) and behavioral capability ( $r=0.440$ ;  $P=.004$ ) for PA over the 4-month

intervention period. An unexpected decrease in exercise self-efficacy for PA was also observed ( $r=-0.364$ ;  $P=.02$ ). No pre- and postintervention changes were observed for outcome expectations ( $r=-.029$ ;  $P=.87$ ), social support from family ( $r=0.103$ ;  $P=.55$ ), or social support from friends ( $r=0.083$ ;  $P=.62$ ).

**Table 4.** Social cognitive theory outcomes.

Variable	Range <sup>a</sup>	Baseline		4 months		P value <sup>b</sup>	Effect size <sup>c</sup>
		Mean (SD)	Median	Mean (SD)	Median		
Outcome expectations	1-5	4.30 (0.89)	4.61	4.41 (0.49)	4.61	.87	-0.029
Self-regulation	1-5	1.96 (0.46)	2.0	2.49 (0.89)	2.40	.01	0.397 <sup>d</sup>
Self-efficacy	1-5	3.95 (0.57)	3.86	3.64 (0.49)	3.63	.02	-0.364 <sup>d</sup>
Social support from friends	8-40	16.53 (9.92)	16.0	18.00 (9.64)	17.0	.62	0.083
Social support from family	10-50	17.25 (7.91)	15.0	17.63 (7.47)	15.0	.55	0.103 <sup>e</sup>
Behavioral capability	1-5	2.30 (1.13)	2.50	3.40 (1.19)	4.0	.004	0.440 <sup>d</sup>

<sup>a</sup>Potential range for each survey measure.

<sup>b</sup>Wilcoxon signed-rank *P* value for changes from baseline to 4 months.

<sup>c</sup>Pearson *r* effect size estimate.

<sup>d</sup>Effect size: medium (0.30-0.49).

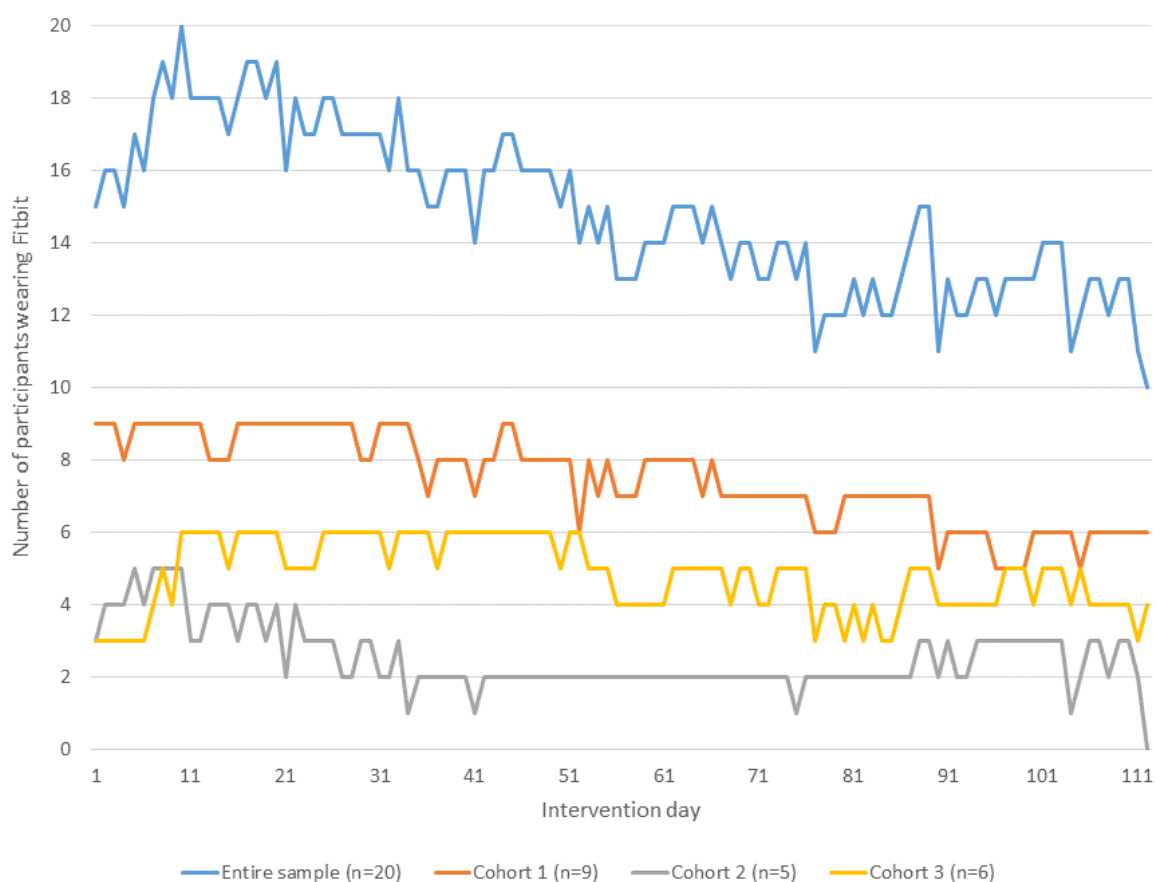
<sup>e</sup>Effect size: small (0.10-0.29).

### Acceptability and Feasibility

#### Fitbit Wear and Activity Tracking Feature

The median number of days participants wore the Fitbit device was 101.2 of the 112-day intervention period (ie, 91% of intervention days). Half of the sample ( $n=10$ ) wore the device

on more than 90% of intervention days, 3 participants wore the Fitbit between 75% and 90% of intervention days, 2 participants wore the device between 50% and 75% of intervention days, and 25% (5/20) of participants wore it for less than 50% of the intervention days. Daily Fitbit wear over time is presented in [Figure 2](#), which shows a trend for decreasing Fitbit wear as the intervention period progressed.

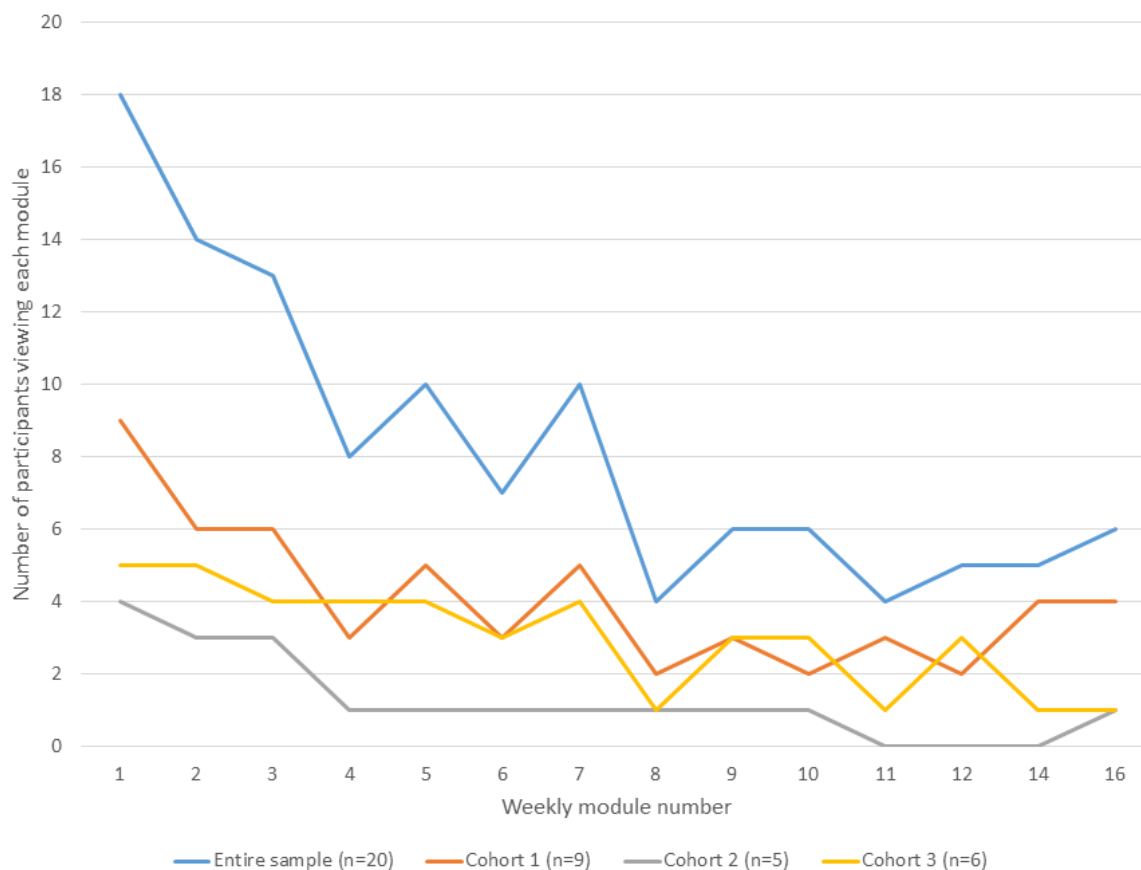
**Figure 2.** Fitbit wear by day and study cohort.

Self-reported feedback regarding the activity tracking feature and the Fitbit activity monitor was generally favorable. Overall, 87% (13/15) of the sample indicated that the combined use of the Fitbit and activity tracking feature available on the app was *very motivating* (n=2), *motivating* (n=7), or *somewhat motivating* (n=4) for increasing PA; 2 participants reported that these features were *not motivating*. The majority of participants also indicated that the activity monitor was *very comfortable* (n=8) or *somewhat comfortable* (n=5) to wear; 1 participant reported that it was *a little comfortable*, and 1 indicated that it was *not comfortable*. With regard to frequency of accessing the activity tracking feature, 47% (7/15) of participants reported viewing the activity tracking feature more than 7 times per week, 7% (1/15) participant reported viewing the tracker 4-6 times per week, 27% (4/15) participants indicated viewing this feature

2-3 times per week, and 20% (3/15) participants reported viewing the feature no more than 1 time per week.

### **Multimedia PA Promotion Modules**

Among the 20 participants enrolled in the study, the median number of weekly multimedia intervention modules viewed was 4 (out of 14 possible). However, among participants who completed the intervention (n=16), the median number of modules viewed was 7. As illustrated in Figure 3, module viewing decreased as the 4-month study progressed. Videos embedded in these multimedia modules were viewed at a lower rate, with only 8 participants pressing the play video on these videos. Among these 8 participants, 3 viewed at least 75% (10/14) of the module videos, 1 viewed 50% (7/14) of the videos, and 4 viewed less than 50% (7/14) of the videos.

**Figure 3.** Module views by week and study cohort.

Participant feedback regarding multimedia modules indicated that 80% (12/15) of the participants found the modules to be *very helpful* (n=3), *helpful* (n=4), or *somewhat helpful* (n=5) for promoting PA. Three participants indicated that they were *not helpful* for promoting PA. In addition, 87% (13/15) of the participants indicated that they were *motivated* (n=4) or *somewhat motivated* (n=9) to be physically active as a result of weekly video and text modules; 2 participants said they were *not motivated*.

### Discussion Board Features

The number of discussion board posts varied by cohort and by week and is shown in Table 5 and graphically displayed in Figure 4. Similar to the trends observed for Fitbit wear and module viewing, discussion board posts decreased as time progressed. Participants reported accessing the message boards at varying frequencies: 20% (4/20) reported viewing the discussion boards 4-6 times per week, 20% (4/20) reported viewing 2-3 times per week, 15% (3/20) reported viewing 1 time per week, and 20% (4/20) reported viewing less than 1 time per week. Self-reported feedback from participants indicated that although the discussion boards were easy to use (12 of 15 participants indicated they did not have any difficulties using the discuss boards), their utility for encouraging PA was somewhat limited, as only 60% (9/15) of participants indicated the discussion boards were *helpful* (n=5) or *somewhat helpful* (n=4) for promoting PA; 40% (6/20) reported they were *not helpful*. Qualitative narratives indicated that participants felt

that the discussion boards lacked interactivity and that additional activities were needed to enhance the discussion board feature and the social support components of the intervention. Quotes provided on the satisfaction survey illustrating the lack of interactivity among participants on the discussion boards included:

*low interaction on the discussion board.*

*lack of camaraderie [on the discussion boards] among participants.*

*lack of interaction [on the discussion boards] with the group.*

Participant narratives describing how we can improve the discussion boards and provisions of social support provided by the intervention included:

*[have a member of the study team] encourage reaching out, connecting and partaking in the discussions.*

*[have] a discussion led by the program.*

*require a check-in [from study staff] 2-3 times per week and setup group exercise activities.*

*[have] a monthly meeting or web chat for participants or perhaps weekly.*

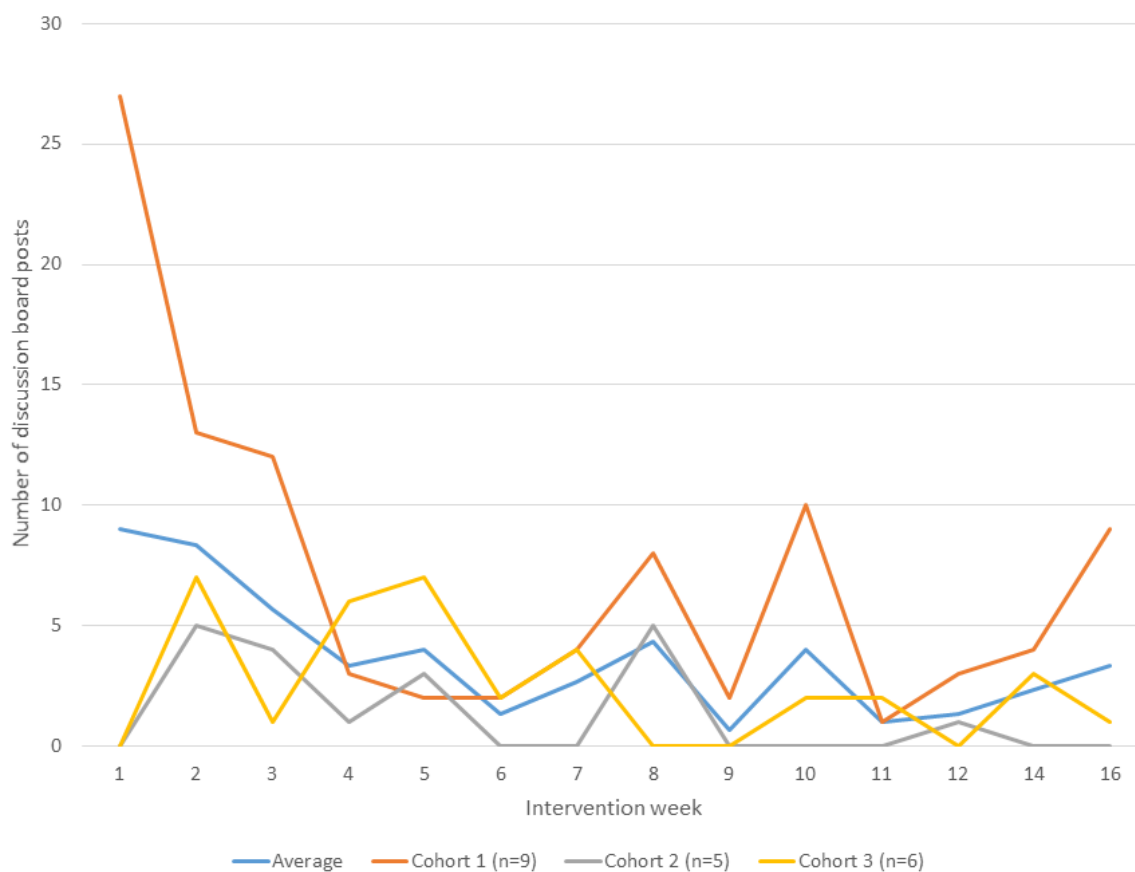
*maybe have [a study team member organize] a meet-up or challenge.*



**Table 5.** Discussion board posts by cohort.

Week	Cohort 1 (n=9)		Cohort 2 (n=5)		Cohort 3 (n=6)	
	Total number participant posts	Frequency of participants posting	Total number participant posts	Frequency of participants posting	Total number participant posts	Frequency of participants posting
1	27	8	0	0	0	0
2	13	5	5	4	7	4
3	12	4	4	2	1	1
4	3	2	1	1	6	4
5	2	2	3	2	7	5
6	2	2	0	0	2	2
7	4	3	0	0	4	4
8	8	3	5	2	0	0
9	2	2	0	0	0	0
10	10	2	0	0	2	2
11	1	1	0	0	2	2
12	3	2	1	1	0	0
13	N/A <sup>a</sup>	N/A	N/A	N/A	N/A	N/A
14	4	3	0	0	3	2
15	N/A	N/A	N/A	N/A	N/A	N/A
16	9	4	0	0	1	1

<sup>a</sup>N/A: not applicable.

**Figure 4.** Discussion board posts by week and study cohort.

### Text Messages

Text message delivery software indicated that text messages were delivered with a high level of fidelity. These data showed that all study text messages were delivered to 19 of the 20 participants. One text message was not delivered to one study participant (ie, sent but received an *undeliverable* response from Twilio). The research team was unable to identify the reason for this message not being deliverable, as the issue self-resolved, and the participant received all other study text messages. Feedback regarding the text messages indicated that 80% (12/15) of participants thought the text messages were *very helpful* (n=2), *helpful* (n=2), or *somewhat helpful* (n=8) for promoting PA.

### Additional Feedback From Participants Regarding Acceptability and Feasibility of the Intervention

Among participants completing the satisfaction survey (n=15), 13 (87%) reported gaining knowledge about PA or exercise from the smartphone-delivered intervention and 14 (93%) indicated that they would recommend the intervention a friend. The one participant indicating she would not recommend the study to a friend stated that the intervention, “didn’t allow for building fitness connections.” This feedback, along with participant suggestions on how we can improve the discussion board features and social support components of the intervention, indicated the need to enhance the intervention to further promote connectedness and *camaraderie* among participants. Specifically, qualitative narratives provided by participants on the satisfaction survey suggested the need for a member of the study team or a PA coach to actively engage with participants and facilitate group-based social support for PA. Despite meaningful feedback provided by participants on how to improve the intervention, many participants expressed appreciation and positive sentiments, supporting the feasibility of the approach. Quotes illustrating this included:

*I gained a lot and will continue to walk more and more.*

*Thank you! I really needed this. I was in a slump.*

*I like that it [the intervention] was interactive.*

*I enjoyed the study. It helped me keep moving.*

*keep all things [referring to the multiple components of the intervention], as everyone is different and [you] never know what motivates a person.*

## Discussion

### Principal Findings

This study evaluated the acceptability and feasibility of a smartphone-delivered PA intervention among African American women. The results indicated that the smartphone-delivered approach increased self-reported PA and enhanced the theoretical mediators of intervention effects, including behavioral capability and self-regulation for PA. Objectively measured metrics of app use and intervention receipt, along with postintervention feedback from participants, highlighted several areas in which the intervention should be refined before larger-scale implementation of the intervention. The results add

to the limited body of research on the use of mHealth PA interventions among midlife to older African American women.

The analyses showed a median increase of approximately 30 minutes of self-reported MVPA per week. Similar increases were also reported for the total weekly estimated energy expenditure. Effect size estimates for self-reported PA were comparable with those for self-reported PA outcomes reported in a previous eHealth and mHealth intervention of similar duration targeting African American women [41] as well as eHealth and mHealth PA interventions among obese women [42]. Evaluation of the SCT constructs targeted by the intervention showed enhancements in behavioral capability and self-regulation. We anticipate that these findings were related to the high level of module viewership during the early weeks of the intervention (ie, weeks 1-3; Figure 3), which largely included content focused on enhancing these psychosocial processes for the successful promotion of PA promotion (Table 1 shows the module content). These findings speak favorably for the intervention increasing knowledge and skill for PA and for participants enacting behavioral strategies to incorporate more PA into daily routine. An unexpected outcome was the decrease in self-efficacy. We speculate that participants may have been overly optimistic regarding their ability to increase PA at baseline, and once they attempted to increase PA, they realized that increasing their activity levels was not as easy as expected. Several participants alluded to this through comments provided on the satisfaction survey, including:

*The program was good. It may have just been a bad time for me as it relates to the time commitment of working out.*

*I wish I had been as active as I inspired[sic] to be.*

*I work full-time and care for my mom after work, so time is precious.*

In addition, although not specifically mentioned by participants in the satisfaction survey, we anticipate that the onset of the SARS-CoV-2 pandemic during the latter stages of intervention delivery for cohorts 2 and 3 also played a role in this outcome. The findings suggest the need to further explore how the intervention can be refined to leverage sources of self-efficacy for the successful promotion of PA.

The lack of changes from pre- to postintervention for outcome expectations may be a result of a ceiling effect, as participants reported relatively high scores on this measure (median scores of 4.61 out of 5, indicating favorable outcome expectations for PA) at both time points. Weak pre- versus postintervention changes in social support from family or friends were unsurprising, considering the lack of engagement in the app’s discussion board feature and qualitative feedback provided by participants. Together, these outcomes highlight a key area for intervention refinement. When originally developed, the purpose of the discussion board feature was to provide a venue for participants to interact with each other and to provide and receive social support. Weekly prompts provided on the discussion boards were included to initiate conversations among participants. However, our data show that these activities were not sufficient to promote social support for PA or engagement among participants on the *Smart Walk* app. Participants desired

synchronous, more frequent, and interactive engagement with the study staff and other study participants. In particular, participants suggested that a member of the research team actively engage with them, provide frequent one-on-one check-ins for support and accountability, and organize group exercise sessions. Given that this is one of the first studies to examine an exclusively technology-mediated PA intervention among previously sedentary midlife to older African American with obesity, the findings provide important insights for researchers developing interventions for this population.

Objectively measured app use provided a great deal of insight regarding the feasibility and acceptability of the smartphone-delivered approach. The relatively high adherence to Fitbit wear, self-reported viewing of the activity tracking feature, and favorable outcomes regarding the motivation provided by these combined features suggest that it is feasible to provide midlife African American women with a Fitbit and allow them to track their activity. Nonetheless, we observed a decrease in Fitbit wear over the course of the study. In future trials, we will incorporate additional strategies to increase adherence to this study component, particularly among participants who demonstrate a decline in wear during the latter stages of intervention delivery. We were also able to deliver text messages to participants with a high degree of fidelity. However, a limitation of the study was that we did not include an item on the satisfaction survey assessing how many of the text messages participants read. Such information would have yielded greater insight into the receipt of the study text messages and will be incorporated into future studies.

Analytic tracking software showed that participants completing the intervention viewed half of the available weekly multimedia PA promotion modules, with a trend for decreased module viewing as the study progressed. Although decreased app use over time is a common issue reported in the mHealth literature [43], these outcomes may indicate the need to revise the modules to make them more interactive and to identify strategies to entice viewership (ie, reminders to view the modules for participants who have not viewed them). Identifying effective strategies to increase module viewing will increase the *dose* of the intervention received by participants, which, in theory, is expected to further enhance PA outcomes. However, this is not definitive. Additional work is needed to identify the optimal amount of intervention receipt needed to promote meaningful changes in PA.

### Limitations and Strengths

Limitations of the study include the small sample size and lack of a control group. Given the purpose of this study was to examine the feasibility of the intervention, rather than efficacy testing, the design was appropriate. Similarly, the multicomponent nature of the intervention does not allow us to draw conclusions as to whether one component is more effective

than another for promoting PA. To answer this type of research question, a multiphase optimization strategy design may be warranted. Other limitations include the use of self-reported PA measures as study outcomes and that the latter stages of intervention delivery for cohorts 2 and 3 coincided with the early stages of the COVID-19 pandemic (February to April 2020). Given that this time was characterized by disruptions of daily routines, uncertainty regarding how the novel coronavirus was transmitted and the likelihood of being infected, and enhanced psychological stress [44,45], these issues likely impacted participants' PA patterns, psychosocial outcomes (particularly self-efficacy for PA), app use, and attrition (3 of the 4 women lost to follow-up were enrolled in these cohorts). Likewise, safety concerns associated with virus transmission inhibited in-person data assessments at 4 months for participants in these 2 cohorts, limiting our ability to examine BMI and blood pressure outcomes among participants. Finally, we originally planned to include wrist-worn accelerometer-measured PA as an outcome measure. However, the collection of these data proved challenging among participants completing their 4-month follow-up assessments during the pandemic. Participants indicated limited interest in wearing the device due to COVID-19-related stress and ongoing disruptions of their daily activities. This resulted in less than half of the participants providing valid wear at the 4-month follow-up. Owing to the high rate of missing accelerometer data, objectively measured PA was not reported here.

Despite these limitations, this study has several strengths. This is one of the few studies to test a mHealth PA intervention among midlife to older African American women. These findings provide important insights into the use of mHealth technology to promote PA among this population at high risk for cardiometabolic diseases. Other strengths include the use of cultural tailoring and behavioral theory in the intervention design. Cultural tailoring is believed to enhance the salience and behavioral outcomes of a behavior change intervention [28,46,47], and numerous studies have shown that theoretically based behavior change interventions are more effective than those that are atheoretical [48-51]. These design considerations address the weaknesses noted in previous reviews of the PA promotion literature among African American women [41,52].

### Conclusions

The results provide preliminary support for the feasibility of the smartphone-based approach to increase PA among insufficiently active midlife to older African American women. However, before larger-scale implementation, several refinements to the intervention are necessary, including enhancing the social support components of the intervention. The findings will be used to inform future research using smartphone-based approaches to promote PA among midlife African American women.

### Acknowledgments

This research was supported by an award from the Arizona State University Institute of Social Science Research.

## Conflicts of Interest

None declared.

## References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity among adults and youth: United States, 2015-2016. NCHS Data Brief. 2017 Oct. URL: <http://www.cdc.gov/nchs/data/databriefs/db288.pdf> [accessed 2020-09-29]
2. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on Epidemiology Prevention Statistics Committee Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2019 Update: a report from the American Heart Association. *Circulation* 2019 Mar 05;139(10):56-528. [doi: [10.1161/CIR.0000000000000659](https://doi.org/10.1161/CIR.0000000000000659)] [Medline: [30700139](https://pubmed.ncbi.nlm.nih.gov/30700139/)]
3. National Diabetes Statistics Report, 2017. Centers for Disease Control and Prevention. Atlanta, GA: U.S. Department of Health and Human Services; 2017. URL: <https://dev.diabetes.org/sites/default/files/2019-06/cdc-statistics-report-2017.pdf> [accessed 2021-03-23]
4. Curtin SC. Trends in cancer and heart disease death rates among adults aged 45-64: United States, 1999-2017. National Vital Statistics Reports. Hyattsville, MD: National Center for Health Statistics; 2019. URL: [https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68\\_05-508.pdf](https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_05-508.pdf) [accessed 2021-03-23]
5. 2008 physical activity guidelines for Americans. Hyattsville, MD: US Department of Health and Human Services; 2008. URL: <https://health.gov/sites/default/files/2019-09/paguide.pdf> [accessed 2021-03-23]
6. Williams WM, Yore MM, Whitt-Glover MC. Estimating physical activity trends among blacks in the United States through examination of four national surveys. *AIMS Public Health* 2018;5(2):144-157 [FREE Full text] [doi: [10.3934/publichealth.2018.2.144](https://doi.org/10.3934/publichealth.2018.2.144)] [Medline: [30094277](https://pubmed.ncbi.nlm.nih.gov/30094277/)]
7. Keller C, Larkey L, Distefano JK, Boehm-Smith E, Records K, Robillard A, et al. Perimenopausal obesity. *J Womens Health (Larchmt)* 2010 May;19(5):987-996. [doi: [10.1089/jwh.2009.1547](https://doi.org/10.1089/jwh.2009.1547)] [Medline: [20380578](https://pubmed.ncbi.nlm.nih.gov/20380578/)]
8. Rosano GM, Vitale C, Marazzi G, Volterrani M. Menopause and cardiovascular disease: the evidence. *Climacteric* 2007 Feb;10 Suppl 1:19-24. [doi: [10.1080/13697130601114917](https://doi.org/10.1080/13697130601114917)] [Medline: [17364594](https://pubmed.ncbi.nlm.nih.gov/17364594/)]
9. Khoudary SR, Thurston RC. Cardiovascular implications of the menopause transition: endogenous sex hormones and vasomotor symptoms. *Obstet Gynecol Clin North Am* 2018 Dec;45(4):641-661. [doi: [10.1016/j.ogc.2018.07.006](https://doi.org/10.1016/j.ogc.2018.07.006)] [Medline: [30401548](https://pubmed.ncbi.nlm.nih.gov/30401548/)]
10. 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:- [FREE Full text] [doi: [10.1177/2055207619839883](https://doi.org/10.1177/2055207619839883)] [Medline: [30944728](https://pubmed.ncbi.nlm.nih.gov/30944728/)]
11. Buchholz SW, Wilbur J, Ingram D, Fogg L. Physical activity text messaging interventions in adults: a systematic review. *Worldviews Evid Based Nurs* 2013 Aug;10(3):163-173. [doi: [10.1111/wvn.12002](https://doi.org/10.1111/wvn.12002)] [Medline: [23746267](https://pubmed.ncbi.nlm.nih.gov/23746267/)]
12. Feter N, 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](https://doi.org/10.1007/s00038-019-01210-7)] [Medline: [30758514](https://pubmed.ncbi.nlm.nih.gov/30758514/)]
13. Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally-relevant Facebook and text message delivered intervention to promote physical activity in African American women: a randomized pilot trial. *BMC Womens Health* 2015 Mar 27;15:30 [FREE Full text] [doi: [10.1186/s12905-015-0186-1](https://doi.org/10.1186/s12905-015-0186-1)] [Medline: [25886945](https://pubmed.ncbi.nlm.nih.gov/25886945/)]
14. Kim BH, Glanz K. Text messaging to motivate walking in older African Americans: a randomized controlled trial. *Am J Prev Med* 2013 Jan;44(1):71-75. [doi: [10.1016/j.amepre.2012.09.050](https://doi.org/10.1016/j.amepre.2012.09.050)] [Medline: [23253653](https://pubmed.ncbi.nlm.nih.gov/23253653/)]
15. McCoy P, Leggett S, Bhuiyan A, Brown D, Frye P, Williams B. Text messaging: an intervention to increase physical activity among african american participants in a faith-based, competitive weight loss program. *Int J Environ Res Public Health* 2017 Mar 29;14(4):- [FREE Full text] [doi: [10.3390/ijerph14040326](https://doi.org/10.3390/ijerph14040326)] [Medline: [28353650](https://pubmed.ncbi.nlm.nih.gov/28353650/)]
16. Zhang J, Jemmott Iii JB. Mobile app-based small-group physical activity intervention for young african american women: a pilot randomized controlled trial. *Prev Sci* 2019 Aug;20(6):863-872. [doi: [10.1007/s11121-019-01006-4](https://doi.org/10.1007/s11121-019-01006-4)] [Medline: [30788692](https://pubmed.ncbi.nlm.nih.gov/30788692/)]
17. Mobile fact sheet. Pew Research Center. 2017. URL: <http://www.pewinternet.org/fact-sheet/mobile/> [accessed 2017-08-31]
18. Joseph RP, Ainsworth BE, Vega-López S, Adams MA, Hollingshead K, Hooker SP, et al. Rationale and design of Smart Walk: a randomized controlled pilot trial of a smartphone-delivered physical activity and cardiometabolic risk reduction intervention for African American women. *Contemp Clin Trials* 2019 Feb;77:46-60 [FREE Full text] [doi: [10.1016/j.cct.2018.12.011](https://doi.org/10.1016/j.cct.2018.12.011)] [Medline: [30576840](https://pubmed.ncbi.nlm.nih.gov/30576840/)]
19. Smith A. Smartphone ownership 2013. Pew Research Center. 2013. URL: <https://www.pewresearch.org/internet/2013/06/05/smartphone-ownership-2013/> [accessed 2020-11-19]
20. Coleman KJ, Ngor E, Reynolds K, Quinn VP, Koebnick C, Young DR, et al. Initial validation of an exercise "vital sign" in electronic medical records. *Med Sci Sports Exerc* 2012;44(11):2071-2076. [doi: [10.1249/MSS.0b013e3182630ec1](https://doi.org/10.1249/MSS.0b013e3182630ec1)] [Medline: [22688832](https://pubmed.ncbi.nlm.nih.gov/22688832/)]

21. 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](#)]
22. Joseph RP, Keller C, Vega-López S, Adams MA, English R, Hollingshead K, et al. A culturally relevant smartphone-delivered physical activity intervention for African American women: development and initial usability tests of smart walk. *JMIR Mhealth Uhealth* 2020 Mar 02;8(3):15346 [FREE Full text] [doi: [10.2196/15346](#)] [Medline: [32130198](#)]
23. Tudor-Locke C, Han H, Aguiar EJ, Barreira TV, Schuna JM, Kang M, et al. How fast is fast enough? Walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. *Br J Sports Med* 2018 Jun;52(12):776-788 [FREE Full text] [doi: [10.1136/bjsports-2017-097628](#)] [Medline: [29858465](#)]
24. Slaght J, Sénéchal M, Hrubeniuk TJ, Mayo A, Bouchard DR. Walking cadence to exercise at moderate intensity for adults: a systematic review. *J Sports Med (Hindawi Publ Corp)* 2017;2017 [FREE Full text] [doi: [10.1155/2017/4641203](#)] [Medline: [28459099](#)]
25. Reese JM, Joseph RP, Cherrington A, Allison J, Kim Y, Spear B, et al. Development of participant-informed text messages to promote physical activity among African American women attending college: a qualitative mixed-methods inquiry. *J Transcult Nurs* 2017 May;28(3):236-242 [FREE Full text] [doi: [10.1177/1043659616644959](#)] [Medline: [27093904](#)]
26. Joseph RP, Ainsworth BE, Mathis L, Hooker SP, Keller C. Utility of social cognitive theory in intervention design for promoting physical activity among African-American women: a qualitative study. *Am J Health Behav* 2017 Sep 01;41(5):518-533 [FREE Full text] [doi: [10.5993/AJHB.41.5.1](#)] [Medline: [28760174](#)]
27. Bandura A. Social foundations of thought and action: a social cognitive framework. Prentice Hall, NJ: Englewood Cliffs; 1986. URL: <https://www.worldcat.org/title/social-foundations-of-thought-and-action-a-social-cognitive-theory/oclc/12080269> [accessed 2021-03-23]
28. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. *Ethn Dis* 1999;9(1):10-21. [Medline: [10355471](#)]
29. Molina L, Sarmiento M, Peñafiel J, Donaire D, Garcia-Aymerich J, Gomez M, et al. Validation of the regicor short physical activity questionnaire for the adult population. *PLoS One* 2017;12(1):e0168148 [FREE Full text] [doi: [10.1371/journal.pone.0168148](#)] [Medline: [28085886](#)]
30. Joseph RP, Keller C, Adams MA, Ainsworth BE. Validity of two brief physical activity questionnaires with accelerometers among African-American women. *Prim Health Care Res Dev* 2016 May;17(3):265-276 [FREE Full text] [doi: [10.1017/S1463423615000390](#)] [Medline: [26178779](#)]
31. Sallis JF, Pinski RB, Grossman RM, Patterson TL, Nader PR. The development of self-efficacy scales for health-related diet and exercise behaviors. *Health Educ Res* 1988;3(3):283-292. [doi: [10.1093/her/3.3.283](#)]
32. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. *Prev Med* 1987 Nov;16(6):825-836. [doi: [10.1016/0091-7435\(87\)90022-3](#)] [Medline: [3432232](#)]
33. Anderson ES, Winett RA, Wojcik JR, Williams DM. Social cognitive mediators of change in a group randomized nutrition and physical activity intervention: social support, self-efficacy, outcome expectations and self-regulation in the guide-to-health trial. *J Health Psychol* 2010 Jan;15(1):21-32. [doi: [10.1177/1359105309342297](#)] [Medline: [20064881](#)]
34. Anderson ES, Wojcik JR, Winett RA, Williams DM. Social-cognitive determinants of physical activity: the influence of social support, self-efficacy, outcome expectations, and self-regulation among participants in a church-based health promotion study. *Health Psychol* 2006 Jul;25(4):510-520. [doi: [10.1037/0278-6133.25.4.510](#)] [Medline: [16846326](#)]
35. Resnick B, Zimmerman SI, Orwig D, Furstenberg AL, Magaziner J. Outcome expectations for exercise scale: utility and psychometrics. *J Gerontol B Psychol Sci Soc Sci* 2000 Nov;55(6):352-356. [doi: [10.1093/geronb/55.6.s352](#)] [Medline: [11078112](#)]
36. Vega WA, Sallis JF, Patterson T, Rupp J, Atkins C, Nader PR. Assessing knowledge of cardiovascular health-related diet and exercise behaviors in Anglo- and Mexican-Americans. *Prev Med* 1987 Sep;16(5):696-709. [doi: [10.1016/0091-7435\(87\)90052-1](#)] [Medline: [3684979](#)]
37. Kay MC, Carroll DD, Carlson SA, Fulton JE. Awareness and knowledge of the 2008 Physical Activity Guidelines for Americans. *J Phys Act Health* 2014 May;11(4):693-698. [doi: [10.1123/jpah.2012-0171](#)] [Medline: [23493071](#)]
38. Hartman SJ, Nelson SH, Weiner LS. Patterns of Fitbit use and activity levels throughout a physical activity intervention: exploratory analysis from a randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Feb 05;6(2):e29 [FREE Full text] [doi: [10.2196/mhealth.8503](#)] [Medline: [29402761](#)]
39. Joseph RP, Pekmezi D, Dutton GR, Cherrington AL, Kim Y, Allison JJ, et al. Results of a culturally adapted internet-enhanced physical activity pilot intervention for overweight and obese young adult African American women. *J Transcult Nurs* 2016 Mar;27(2):136-146 [FREE Full text] [doi: [10.1177/1043659614539176](#)] [Medline: [24934566](#)]
40. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](#)] [Medline: [16204405](#)]
41. Joseph RP, Roysse KE, Benitez TJ. A systematic review of electronic and mobile health (e- and mHealth) physical activity interventions for African American and Hispanic women. *J Phys Act Health* 2019 Mar 01;16(3):230-239. [doi: [10.1123/jpah.2018-0103](#)] [Medline: [30782040](#)]

42. Cotie LM, Prince SA, Elliott CG, Ziss MC, McDonnell LA, Mullen KA, et al. The effectiveness of eHealth interventions on physical activity and measures of obesity among working-age women: a systematic review and meta-analysis. *Obes Rev* 2018 Oct;19(10):1340-1358. [doi: [10.1111/obr.12700](https://doi.org/10.1111/obr.12700)] [Medline: [30156044](https://pubmed.ncbi.nlm.nih.gov/30156044/)]
43. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. *J Exerc Rehabil* 2017 Feb;13(1):3-11 [FREE Full text] [doi: [10.12965/jer.1732928.464](https://doi.org/10.12965/jer.1732928.464)] [Medline: [28349027](https://pubmed.ncbi.nlm.nih.gov/28349027/)]
44. Tull MT, Edmonds KA, Scamaldo KM, Richmond JR, Rose JP, Gratz KL. Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. *Psychiatry Res* 2020 Jul;289:113098 [FREE Full text] [doi: [10.1016/j.psychres.2020.113098](https://doi.org/10.1016/j.psychres.2020.113098)] [Medline: [32434092](https://pubmed.ncbi.nlm.nih.gov/32434092/)]
45. Park CL, Russell BS, Fendrich M, Finkelstein-Fox L, Hutchison M, Becker J. Americans' COVID-19 stress, coping, and adherence to CDC guidelines. *J Gen Intern Med* 2020 Aug 29;35(8):2296-2303 [FREE Full text] [doi: [10.1007/s11606-020-05898-9](https://doi.org/10.1007/s11606-020-05898-9)] [Medline: [32472486](https://pubmed.ncbi.nlm.nih.gov/32472486/)]
46. Resnicow K, Soler R, Braithwaite RL, Ahluwalia JS, Butler J. Cultural sensitivity in substance use prevention. *J Community Psychol* 2000;28(3):271-290. [doi: [10.1002/\(sici\)1520-6629\(200005\)28:3<271::aid-jcop4>3.0.co;2-i](https://doi.org/10.1002/(sici)1520-6629(200005)28:3<271::aid-jcop4>3.0.co;2-i)]
47. Resnicow K, Braithwaite R, Dilorio CK. Applying theory to culturally diverse and unique populations. In: Glanz K, Lewis F, Rimer B, editors. *Health Behavior and Health Education: Theory, Research and Practice*. San Francisco, CA: Jossey-Bass; 2002:485-509.
48. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. *Annu Rev Public Health* 2010;31:399-418. [doi: [10.1146/annurev.publhealth.012809.103604](https://doi.org/10.1146/annurev.publhealth.012809.103604)] [Medline: [20070207](https://pubmed.ncbi.nlm.nih.gov/20070207/)]
49. Levy RL, Finch EA, Crowell MD, Talley NJ, Jeffery RW. Behavioral intervention for the treatment of obesity: strategies and effectiveness data. *Am J Gastroenterol* 2007;102(10):2314-2321. [doi: [10.1111/j.1572-0241.2007.01342.x](https://doi.org/10.1111/j.1572-0241.2007.01342.x)] [Medline: [17561967](https://pubmed.ncbi.nlm.nih.gov/17561967/)]
50. 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](https://doi.org/10.1186/1479-5868-9-52)] [Medline: [22546283](https://pubmed.ncbi.nlm.nih.gov/22546283/)]
51. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](https://doi.org/10.1037/0033-2909.133.4.673)] [Medline: [17592961](https://pubmed.ncbi.nlm.nih.gov/17592961/)]
52. Whitt-Glover MC, Keith NR, Ceaser TG, Virgil K, Ledford L, Hasson RE. A systematic review of physical activity interventions among African American adults: evidence from 2009 to 2013. *Obes Rev* 2014 Oct;15 Suppl 4:125-145. [doi: [10.1111/obr.12205](https://doi.org/10.1111/obr.12205)] [Medline: [25196410](https://pubmed.ncbi.nlm.nih.gov/25196410/)]

## Abbreviations

**MET:** metabolic equivalent

**mHealth:** mobile health

**MVPA:** moderate- to vigorous-intensity physical activity

**PA:** physical activity

**SCT:** social cognitive theory

*Edited by L Buis; submitted 26.01.21; peer-reviewed by B Staffileno, S Nyenhuis; comments to author 09.02.21; revised version received 05.03.21; accepted 09.03.21; published 22.04.21.*

*Please cite as:*

*Joseph RP, Ainsworth BE, Hollingshead K, Todd M, Keller C*

*Results of a Culturally Tailored Smartphone-Delivered Physical Activity Intervention Among Midlife African American Women: Feasibility Trial*

*JMIR Mhealth Uhealth* 2021;9(4):e27383

URL: <https://mhealth.jmir.org/2021/4/e27383>

doi: [10.2196/27383](https://doi.org/10.2196/27383)

PMID: [33885368](https://pubmed.ncbi.nlm.nih.gov/33885368/)

©Rodney P Joseph, Barbara E Ainsworth, Kevin Hollingshead, Michael Todd, Colleen Keller. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 22.04.2021. 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.

Original Paper

# Evaluation of Mood Check-in Feature for Participation in Meditation Mobile App Users: Retrospective Longitudinal Analysis

Jennifer Huberty<sup>1</sup>, PhD; Jeni Green<sup>1</sup>, PhD; Megan Puzia<sup>2</sup>, MS; Chad Stecher<sup>1</sup>, PhD

<sup>1</sup>College of Health Solutions, Arizona State University, Phoenix, AZ, United States

<sup>2</sup>Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

**Corresponding Author:**

Jeni Green, PhD

College of Health Solutions

Arizona State University

500 N 3rd St

Phoenix, AZ, 85004

United States

Phone: 1 8019278558

Email: [jeni.green@asu.edu](mailto:jeni.green@asu.edu)

## Abstract

**Background:** Mindfulness meditation smartphone apps may improve mental health but lack evidence-based behavioral strategies to encourage their regular use for attaining mental health benefits. In October 2019, the Calm mindfulness meditation app introduced a mood check-in feature, but its effects on participation in meditation have yet to be tested.

**Objective:** The objective of this study was to investigate how a mood check-in feature impacts meditation behavior in Calm app subscribers.

**Methods:** This was a retrospective longitudinal analysis of mobile app usage data from a random sample of first-time subscribers to the Calm app (n=2600) who joined in summer 2018 or summer 2019. The mood check-in feature allows users to rate their mood using an emoji after completing a meditation session and displays a monthly calendar of their past mood check-ins. Regression analyses were used to compare the rate of change in meditation behavior before and after the introduction of mood check-ins and to estimate how usage of mood check-ins was associated with individuals' future meditation behavior (ie, intent-to-treat effects). Additional regression models examined the heterogenous effect of mood check-ins between subscribers who were active or inactive users prior to the introduction to mood check-ins (ie, above or below the median number of weeks with any meditation within their cohort). In order to confirm the specific associations between mood check-ins and meditation engagement, we modeled the direct relationship between the use of mood check-ins in previous weeks and subsequent meditation behavior (ie, treatment on the treated effects).

**Results:** During the first 9 months of their subscription, the 2019 cohort completed an average of 0.482 more sessions per week (95% CI 0.309 to 0.655) than the 2018 cohort; however, across both cohorts, average weekly meditation declined (-0.033 sessions per week, 95% CI -0.035 to -0.031). Controlled for trends in meditation before mood check-ins and aggregate differences between the 2018 and 2019 samples, the time trend in the number of weekly meditation sessions increased by 0.045 sessions among the 2019 cohort after the introduction of mood check-ins (95% CI 0.039 to 0.052). This increase in meditation was most pronounced among the inactive subscribers (0.063 sessions, 95% CI 0.052 to 0.074). When controlled for past-week meditation, use of mood check-ins during the previous week was positively associated with the likelihood of meditating the following week (odds ratio 1.132, 95% CI 1.059 to 1.211); however, these associations were not sustained beyond 1 week.

**Conclusions:** Using mood check-ins increases meditation participation in Calm app subscribers and may be especially beneficial for inactive subscribers. Mobile apps should consider incorporating mood check-ins to help better engage a wider range of users in app-based meditation, but more research is warranted.

(*JMIR Mhealth Uhealth* 2021;9(4):e27106) doi:[10.2196/27106](https://doi.org/10.2196/27106)

**KEYWORDS**

adherence; meditation; mindfulness; mood; smartphone application; app; engagement; mHealth; mental health; behavior

## Introduction

There are an estimated 20,000 mental health smartphone apps currently available for download [1]. Mindfulness meditation apps are a popular type of mental health app, with over 290 available for download [1]. To date, Headspace and Calm are the leading mindfulness meditation apps with 65 and 100 million total downloads, respectively [2,3]. Interventions using mindfulness meditation apps indicate small- to medium-sized effects on improvements in depression and anxiety, life satisfaction, and positive affect [4], and greater engagement with mental health apps is associated with larger reductions in mental health symptoms [5]. Thus, mindfulness meditation apps are easily accessible, feasible, and cost-effective tools for promoting mental health and well-being on a large scale [4]. Despite the potential benefits of using these apps, not everyone remains engaged [5], and existing mobile Health strategies to increase mindfulness meditation app engagement have yet to identify techniques that can increase participation, both among subscribers who are active users (ie, currently using an app on a regular or semiregular basis) and those who are inactive users (ie, those who subscribed to an app but did not or no longer use it).

Lack of engagement is a well-documented barrier confronting mobile app interventions. A recent systematic review and meta-analysis [5] reported that participation in mental health app-based interventions consistently decreased over time, and percentages of participants who adhere to all prescribed sessions vary widely between studies and found that, in studies targeting depression, anxiety, or stress, average intervention adherence rates were 34%, 36%, and 41%, respectively. Adherence reported in studies [6,7] testing the Calm meditation app were similar. For example, in an 8-week intervention to improve stress in college students, only 22% of participants adhered to the prescribed meditation schedule (10 minutes per day) through the completion of the intervention [6]. In a study [7] to improve cancer-related symptoms (eg, fatigue, pain, anxiety, sleep disturbance) in patients with blood cancer, meditation participation decreased from an average of 86 minutes per week to 53 minutes per week over the course of the 4-week intervention period (a 38% decline). In both studies [6,7], adherence declined despite high initial motivation and high rates of satisfaction with the meditation app and perceived benefits of using apps to meditate. In real-world settings (eg, noninterventions), rates of app usage also decline rapidly; 71% of users stop using commercially available apps within 3 months of starting a subscription despite having sufficient initial interest and motivation to purchase an app subscription (which range from US \$20 to \$70) [8,9]. There is a limited understanding of how to re-engage meditation app subscribers after long periods of inactivity or after stopping app use entirely [10]. Identifying strategies to increase mobile app use among subscribers who are active or inactive is particularly important for developers of mindfulness meditation apps that are associated with large improvements in mental health only among subscribers who are highly active users [5].

Self-tracking app usage is a common component of existing mental health apps [11]. In general, apps that utilize behavioral

self-tracking features have less usage reduction and a lower likelihood of abandonment compared to apps without these features [12]. Mood check-ins (ie, assessing and documenting one's mood after meditation) may be a particularly useful self-tracking strategy to improve engagement in subscribers who are active users or re-engage subscribers who are inactive users of app-based meditation for several reasons. First, tracking one's mood may help users better recognize the immediate benefits of meditation. Second, mood check-ins may increase emotional self-awareness, which may further improve mental health symptoms [11]. Third, mood check-ins align with several constructs from social cognitive theory, which suggests that mood check-ins may encourage behavior change by increasing positive reinforcement for meditation, self-efficacy, and self-regulation (ie, self-monitoring, judgment, affective reaction) [13,14]. Finally, as meditation requires continued practice in order to develop mindfulness skills and attain the corresponding mental health benefits, meditation may not always provide clear and immediate feedback [15]; therefore, mood check-ins may provide tangible feedback to the user regarding their mental health in the moment and over time. This feedback increases the salience of both the immediate and future benefits of meditation, and thus, may increase meditation engagement by combatting myopic loss aversion, a decision-making bias commonly noted by behavioral economists, in which short-term preferences lead people to undervalue health behaviors (eg, meditation) that produce future benefits [16]. Mood check-ins, therefore, have a strong theoretical justification for improving app engagement; however, their effect on active and inactive subscribers' app-based meditation practices has not yet been evaluated.

In October 2019, the Calm mindfulness meditation app introduced a mood check-in feature that is based on constructs of social cognitive theory [13,14]. Therefore, the purpose of this study was to use app usage data from Calm subscribers to determine whether the introduction of mood check-ins within the Calm app increased meditation engagement.

## Methods

### Overview

We used an event-study design (or difference-in-differences model) to measure the change in meditation use after the introduction of mood check-ins among a random sample of users who had subscribed 3 months before mood check-ins were introduced, and compared that difference to the change in meditation use among a random sample of users who subscribed in the same month the year prior (ie, intent-to-treat effects). We also used this framework to examine whether mood check-ins had a heterogeneous effect on meditation engagement across subscribers in our sample who were active and inactive users prior to the introduction of mood check-ins (ie, above or below the median number of weeks with any meditation within their cohort). To confirm that it was specifically mood check-ins that increased engagement, we modeled the direct relationship between use of mood check-ins during the previous weeks and subsequent meditation behavior (ie, treatment on the treated effects). These findings may help to design future behavioral



interventions for promoting mindfulness meditation app engagement and thus improving subscribers' mental health outcomes.

### Ethics

This was a retrospective longitudinal analysis of mobile app usage data and was approved by the Arizona State University institutional review board (STUDY 00011292). Consent for use of data was provided by subscribers who agreed to the Calm privacy policy, which states that user information may be used for research and shared with third parties in anonymous or aggregate form.

### Participants and Data Source

Usage data were compiled from a random sample of 2600 first-time Calm app subscribers who joined in June or July 2018 (n=1300) or in June or July 2019 (n=1300). Subscribers were eligible for selection based only on the date of their initial paid subscription to the Calm app and were chosen using a randomization function in SQL. The Calm app introduced the mood check-in feature in October 2019; therefore, users who joined in 2018 had not been exposed to mood check-ins over the first 15 to 16 months of their subscription (ie, control group), whereas those who subscribed in 2019 were introduced to mood check-ins between 3 and 4 months after subscribing. No other major changes were introduced in the app during these time periods. Data for both groups were restricted to the first 9 months of app use to avoid analyzing usage patterns during the COVID-19 pandemic (February 2020) among the group that joined in 2019.

### The Calm App

The Calm app is a meditation app that is commercially available. Calm provides general guided meditation (eg, 10-minute Daily Calm, various individual and series meditations, and sleep-specific meditations) grounded in mindfulness-based stress reduction [17] and Vipassana meditation [18] and Sleep Stories, grounded in sensory immersion and present moment awareness. The Daily Calm meditations are 10 minutes in length and include topics that change daily targeted for both the beginner and advanced meditator (eg, karma, distraction, self-compassion). The individual and series meditations include 3- to 30-minute meditations with topics such as stress, self-care, and anxiety (and include meditations for beginners). The series meditations are designed to be practiced daily for 7 days and include topics such as "7 Days of Soothing Pain," "7 Days of Managing Stress," or "7 Days of Calming Anxiety." The sleep-specific meditations are intended to soothe the body and mind into sleep through deep, progressive relaxation. Finally, the Sleep Stories are designed to help users fall asleep while listening to a variety of stories including fiction or nature-based.

Calm recently added a mood check-in feature which allows participants to rate their mood using an emoji after completing a meditation session; users then receive preset feedback messages related to their mood selection (eg, let negative thoughts be, let them go) and a monthly calendar of past mood check-ins that is displayed after checking-in. This feature is consistent with behavioral strategies (eg, reinforcement,

self-monitoring [13,14]) that may help individuals participate in and maintain health behaviors.

### Statistical Analyses

#### *Cohort Effects of the Introduction of the Mood Check-In Feature (Intent to Treat)*

App usage data (number of meditation sessions completed, total minutes of meditation with the app, and likelihood of completing any meditation session) were aggregated by week over the first 9 months of users' subscriptions. To estimate the overall effect of mood check-ins on meditation behavior over time, we used linear regression analyses. We compared the rate of change in meditation behavior measures in the 2019 cohort before and after the introduction of mood check-ins with those of the 2018 cohort during the corresponding time period in the previous year. We used a difference-in-differences approach, in which estimations of changes in meditation behavior were allowed to differ linearly for the periods before and after the mood check-in feature was introduced to determine the intent-to-treat effect. Outcomes were regressed on to a linear measure of weeks since starting to use the app, an indicator variable equal to 1 for the 2019 sample, and an indicator variable equal to 1 for the weeks after the mood check-in feature became available in 2019 (evaluated for the same week in both the 2018 and 2019 cohort), as well as all 2-way and 3-way interactions between the three variables. The model coefficient of interest was the estimated parameter for the change in weekly time trend after mood check-ins became available to the 2019 cohort—the variable *weeks×2019 cohort×mood check-ins*. All analyses were conducted using Stata software (version 16.1; StataCorp LLC).

To investigate whether the effects of mood check-ins were similar for active versus inactive users, we created separate models for active and inactive subscribers. Active and inactive status was defined by the number of weeks in which a subscriber completed at least 1 meditation session during the first 17 weeks of their subscription (corresponding to the period before mood check-ins were introduced in 2019). Users were classified as active subscribers if they were above the median number of weeks with any meditation within their cohort or as inactive subscribers if they were below the median number of meditation weeks threshold within their cohort. Thresholds were 2 and 3 weeks, respectively, for the 2018 and 2019 cohorts.

#### *Association Between Using Mood Check-ins and Future Meditation Behavior (Treatment on the Treated)*

To confirm the specific direct relationship between using mood check-ins and meditation engagement (treatment on the treated effect), we examined associations between an individual's use of mood check-ins during the previous week and that person's future meditation behavior by creating a second set of linear models regressing (1) the likelihood of any meditation during the week and (2) the weekly number of meditation sessions on the number of mood check-ins during the prior week (ie, a 1-unit lag). For these analyses, weekly app usage data were assessed for a 16-week period—8 weeks prior to and following the release of the mood check-in feature in October 2019. We created models for the entire 2019 cohort, active subscribers, and inactive subscribers. To control for differences in aggregate

meditation behavior between subscribers who did and did not use the mood check-in feature, the models controlled for estimated meditation behavior during the prior week. To examine the duration of associations between using mood check-ins and future meditation behavior, we created models predicting subsequent weekly meditation using mood check-ins and meditation behavior measured over the previous 7 weeks.

## Results

### Introduction of the Mood Check-in Feature (Intent-to-Treat Effect)

Summary statistics for weekly meditation practices using the Calm app between the 2018 and 2019 cohorts show significantly greater meditation usage among those who joined in 2019 versus 2018 (Table 1). Compared to users who joined Calm in the summer of 2018, those who joined in the summer of 2019 were 0.122 percentage points more likely to meditate with the Calm app in a given week ( $P < .001$ ), meditated for an average of 0.587 more sessions per week ( $P < .001$ ), and meditated for

approximately 7.761 more minutes per week ( $P < .001$ ). Similar patterns were observed among the active and inactive subscribers (Table 1).

The difference-in-differences model (Figure 1) shows that there were negative linear time trends in weekly meditation sessions occurring at an equal rate in both cohorts prior to the introduction of mood check-ins. This equivalence in trends prior to the introduction of the mood check-in feature for the 2019 cohort satisfies the parallel trends assumption (see Figure S1, Multimedia Appendix 1 for trends among active and inactive subscribers). An increase in meditation sessions occurred following the introduction of the mood check-in feature (Model 1, Table 2): across both cohorts, average weekly meditation sessions declined ( $-0.033$  sessions per week, 95% CI  $-0.035$  to  $-0.031$ ), and in general, the 2019 cohort completed an average of 0.482 more sessions per week (95% CI 0.309 to 0.655) than the 2018 cohort completed. Importantly, the weekly change in meditation sessions significantly increased by 0.045 sessions per week (95% CI 0.039 to 0.052) in the 2019 cohort after the introduction of the mood check-in feature.

**Table 1.** Calm app usage between 2018 and 2019 cohorts during the full study period.

Subscriber meditation characteristics	2018 cohort <sup>a</sup> (n=1300), mean (SD)	2019 cohort <sup>b</sup> (n=1300), mean (SD)	Difference <sup>c</sup>	P value
<b>All subscribers<sup>d</sup></b>				
Any weekly meditation <sup>e</sup>	0.22 (0.42)	0.34 (0.48)	-0.122	<.001
Weekly meditation sessions	0.87 (2.49)	1.45 (3.28)	-0.587	<.001
Weekly meditation minutes	11.03 (35.54)	18.79 (46.14)	-7.761	<.001
<b>Active subscribers<sup>f</sup></b>				
Any weekly meditation	0.39 (0.49)	0.57 (0.50)	-0.178	<.001
Weekly meditation sessions	1.56 (3.19)	2.51 (4.05)	-0.957	<.001
Weekly meditation minutes	19.83 (46.12)	32.48 (57.69)	-12.648	<.001
<b>Inactive subscribers<sup>g</sup></b>				
Any weekly meditation	0.02 (0.14)	0.07 (0.26)	-0.052	<.001
Weekly meditation sessions	0.03 (0.30)	0.16 (0.94)	-0.129	<.001
Weekly meditation minutes	0.40 (3.77)	2.11 (12.72)	-1.708	<.001

<sup>a</sup>Those who subscribed to the Calm app in June or July of 2018.

<sup>b</sup>Those who subscribed in either June or July of 2019.

<sup>c</sup>Difference between 2018 and 2019 cohort weekly averages.

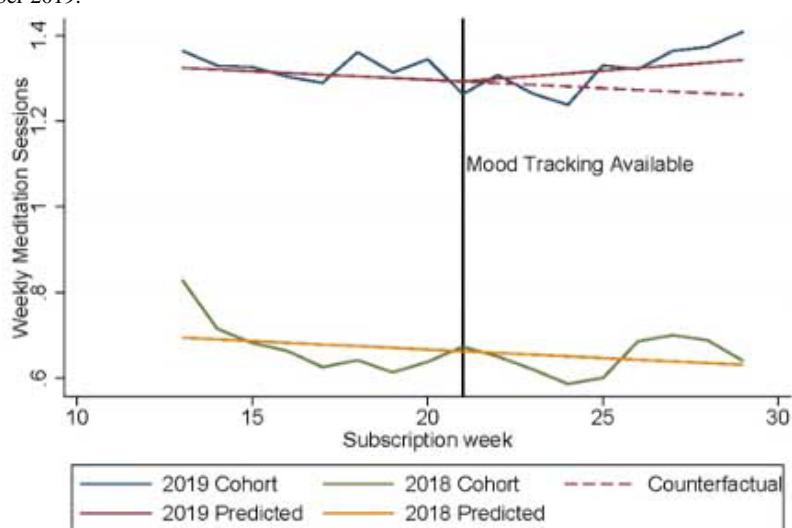
<sup>d</sup>39,000 and 39,000 observations for the 2018 and 2019 cohorts, respectively.

<sup>e</sup>Any weekly meditation describes the estimated likelihood that a meditation session was completed during a given week.

<sup>f</sup>21,330 and 21,420 observations for the 2018 and 2019 cohorts, respectively.

<sup>g</sup>17,670 and 17,580 observations for the 2018 and 2019 cohorts, respectively.

**Figure 1.** Average weekly meditation sessions for the 2018 and 2019 cohorts along with the estimated trend before and after the introduction of the mood check-in feature in October 2019.



Because the trends in weekly meditation sessions are not accurately described by a linear time trend over the study period, we focused our analyses on the 8-week periods before and after mood check-ins were introduced to the 2019 cohort and the same time period in the 2018 cohort (Model 2). The estimated effect of mood check-ins on weekly meditation sessions is attenuated but remains positive and statistically significant (Table 2). Specifically, the time trend in the number of weekly meditation sessions significantly increased by 0.012 over this period (95% CI 0.005 to 0.018). Based on this increased time trend, the cumulative effect of 8 weeks of mood check-ins was

an estimated increase of approximately 1 meditation session. Similar time trends and statistically significant differences in app usage were observed when estimating the effect on weekly meditation minutes, resulting in an estimated increase of 5.1 minutes over the 8 weeks of mood check-ins (Table S1 and Figure S2-S3, Multimedia Appendix 1).

Table 2 shows that the primary effect of mood check-ins on meditation behavior was experienced by inactive subscribers. Specifically, the cumulative effect of 8 weeks of mood check-ins was an estimated increase of 1.7 weekly sessions for the inactive subscribers.

**Table 2.** Effect (ordinary least squares estimates) of the introduction of the mood check-in feature on average weekly meditation sessions.

Subscriber variables	Model 1: full study period		Model 2: 8 weeks before and after	
	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value
<b>All subscribers<sup>a</sup></b>				
Subscription week	-0.033 (-0.035, -0.031)	<.001	-0.006 (-0.009, -0.003)	<.001
2019 cohort	0.482 (0.309, 0.655)	<.001	0.623 (0.455, 0.791)	<.001
Week×2019 cohort×mood check-ins	0.045 (0.039, 0.052)	<.001	0.012 (0.005, 0.018)	<.001
<b>Active subscribers<sup>b</sup></b>				
Subscription week	-0.060 (-0.063, -0.056)	<.001	-0.014 (-0.019, -0.008)	<.001
2019 cohort	0.811 (0.528, 1.094)	<.001	1.061 (0.781, 1.341)	<.001
Week×2019 cohort×mood check-ins	0.063 (0.052, 0.074)	<.001	0.004 (-0.007, 0.015)	.46
<b>Inactive subscribers<sup>c</sup></b>				
Subscription week	-0.001 (-0.002, 0.000)	.20	0.003 (0.001, 0.005)	.003
2019 cohort	0.070 (0.034, 0.107)	<.001	0.082 (0.026, 0.138)	.004
Week×2019 cohort×mood check-ins	0.026 (0.023, 0.029)	<.001	0.021 (0.017, 0.026)	<.001

<sup>a</sup>78,000 and 49,400 observations for model 1 and model 2, respectively.

<sup>b</sup>42,750 and 27,075 observations for model 1 and model 2, respectively.

<sup>c</sup>32,250 and 22,325 observations for model 1 and model 2, respectively.

### Association Between Using Mood Check-Ins and Future Meditation Behavior (Treatment on the Treated)

Table 3 presents the average weekly meditation behavior

between those among the 2019 cohort who did and did not use the mood check-in feature. Compared to nonusers, subscribers who used the feature were more likely to meditate, completed more meditation sessions, and spent more time meditating in a given week.

**Table 3.** Meditation behavior of users of mood check-ins vs nonusers of mood check-ins.

Subscriber meditation characteristics	Nonusers, mean (SD)	Users, mean (SD)	Difference <sup>a</sup>	P value
<b>All subscribers<sup>b</sup></b>				
Likelihood of any weekly meditation <sup>c</sup>	0.21 (0.40)	0.37 (0.48)	-0.122	<.001
Weekly meditation sessions	0.75 (2.93)	1.66 (3.25)	-0.587	<.001
Weekly meditation minutes	10.37 (44.48)	20.91 (44.85)	-7.761	<.001
<b>Active subscribers<sup>d</sup></b>				
Likelihood of any weekly meditation	0.42 (0.49)	0.52 (0.50)	-0.178	<.001
Weekly meditation sessions	1.55 (4.04)	2.45 (3.78)	-0.957	<.001
Weekly meditation minutes	21.31 (62.48)	30.97 (52.94)	-12.648	<.001
<b>Inactive subscribers<sup>e</sup></b>				
Likelihood of any weekly meditation	0.05 (0.22)	0.13 (0.34)	-0.052	<.001
Weekly meditation sessions	0.16 (1.48)	0.36 (1.32)	-0.129	<.001
Weekly meditation minutes	2.43 (20.78)	4.41 (16.67)	-1.708	<.001

<sup>a</sup>Difference between nonuser and user weekly averages.

<sup>b</sup>4660 and 8340 observations for nonusers and users, respectively.

<sup>c</sup>Any weekly meditation describes the likelihood that a meditation session was completed during a given week.

<sup>d</sup>1960 and 5180 observations for nonusers and users, respectively.

<sup>e</sup>2700 and 3160 observations for nonusers and users, respectively.

To examine the association between using mood check-ins and future meditation behavior, we estimated the relationship between using mood check-ins during the prior week and the likelihood of any meditation during the subsequent week for the 8 weeks after mood check-ins were introduced. The results are presented in Table 4. All models control for the aggregate

trends in the likelihood of any weekly meditation both before and after the introduction of mood check-ins, as well as an indicator for any weekly meditation during the previous week. Using the mood check-ins in the previous week increased the odds of meditating by 1.132 (95% CI 1.059 to 1.211). This pattern was similar across active and inactive subscribers.

**Table 4.** Effect (logistic panel regression estimates) of past mood check-ins on the likelihood of weekly meditation.

Predictor	All Subscribers (n=75,400)		Active subscribers (n=41,325)		Inactive subscribers (n=34,075)	
	OR <sup>a</sup> (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Subscription week	0.963 (0.959, 0.966)	<.001	0.957 (0.953, 0.961)	<.001	1.005 (0.995, 1.015)	.33
2019 cohort	2.286 (1.949, 2.681)	<.001	2.055 (1.813, 2.330)	<.001	3.269 (2.697, 3.963)	<.001
Week×2019 cohort×mood check-ins	1.041 (1.030, 1.053)	<.001	1.024 (1.012, 1.037)	<.001	1.047 (1.022, 1.073)	<.001
Lagged 1-week meditation minutes	1.040 (1.038, 1.041)	<.001	1.038 (1.037, 1.040)	<.001	1.043 (1.038, 1.047)	<.001
Lagged 1-week total mood check-ins	1.132 (1.059, 1.211)	<.001	1.127 (1.047, 1.213)	.001	1.147 (1.011, 1.301)	.03

<sup>a</sup>OR: odds ratio.

To examine the long-term relationship between past mood check-ins and subsequent meditation behavior, we estimated associations between 7 weeks of lagged mood check-ins and future meditation engagement (Table S3, Multimedia Appendix 1). Each model also controlled for aggregate trends in the likelihood of any weekly meditation among each cohort over the sample period. The results showed that any past meditation was significantly associated with greater odds of meditation in

a given week, but importantly, the magnitude of this positive relationship declined for meditation that occurred further in the past. The predictive relationship between using the mood check-in feature and the odds of future meditation was not sustained beyond 1 week.

## Discussion

To our knowledge, this is the first study to examine if the introduction of a mood check-in feature within a mindfulness meditation app increased meditation engagement. Our findings demonstrate that users were more likely to participate in meditation after mood check-ins were available (ie, 0.045 more weekly sessions 2019 subscribers following the introduction of mood check-ins compared to 2018 subscribers during the corresponding time period) and that this effect was stronger for inactive subscribers than for active subscribers. Additionally, we found that use of mood check-ins specifically increased the likelihood of future meditation engagement, even when controlling for past meditation behavior and aggregate trends in meditation patterns over time. Use of mood check-ins during the previous week was positively associated with likelihood of meditating the following week; however, use of mood check-ins beyond 1 week prior was not predictive of future meditation.

The 2019 cohort had an estimated cumulative increase in meditation of approximately 1 session (or 5 minutes), 8 weeks after the introduction of mood check-ins, compared to that of the 2018 cohort. Additionally, subscribers who used mood check-ins (compared to those who did not use mood check-ins) were more likely to meditate, completed more meditation sessions, and spent more time meditating on a given week. Importantly, these associations were strongest among previously inactive users. These data highlight the potential for mood check-ins to help subscribers participate in meditation after a period of low engagement. Existing literature suggests mood self-tracking helps app users increase their awareness of their mood patterns and, as such, helps with self-regulation and emotional well-being [19]. The mood check-in feature was also designed to increase the salience of immediate meditation benefits. The mood check-in calendar display may help users better visualize the long-term benefits of meditation and perhaps maintain meditation behavior by improving self-efficacy or even acting as a behavioral reinforcement [14]. Additionally, the increased salience of future long-term mental health benefits may help users overcome present-biased time preferences, allowing users to better evaluate the merits of mindfulness meditation and thus increase their meditation practice. However, to the authors' knowledge, no studies have disentangled the potential mechanisms through which mood check-ins impact the use of a consumer-based meditation app. Future research should further examine how meditation and health outcomes are impacted by mood check-ins and investigate the behavioral channels that underlie these effects. Additional studies are also needed to investigate the use of mood check-ins over longer periods of time to see if and how improvements in meditation habits are maintained.

The effect of mood check-ins was greater for inactive subscribers (ie, cumulative effect of 8 weeks of mood check-ins increased their meditation practice by an estimated 1.7 sessions) than it was for active subscribers (ie, practices did not change). To our knowledge, there is no research evaluating app-based meditation behaviors and factors that influence continued engagement. According to the Transtheoretical Model and the Stages of Change, there is a distinction between people in the

precontemplation or contemplation stages and those in the action or maintenance stages [20]. Those in the action stage are actively involved in taking steps to practice a health behavior, while those in the maintenance stages have ongoing practice of a health behavior and are less likely to relapse [20]. In our data, inactive subscribers may be those in the precontemplation or contemplation stages and may benefit from reinforcement management (such as mood check-ins) to help them move into the action stage, while active subscribers may already be in the maintenance stage, working toward sustaining a behavior rather than increasing their efforts. Interventions using mobile apps may consider evaluating stages of change and incorporating specific reinforcement strategies, such as mood check-ins, to move participants from the action to the maintenance stage.

We were not able to determine whether the benefits of mood check-ins come from individuals consciously deciding to meditate based on their mood or if this effect occurs on a subconscious level. If occurring subconsciously, the meditation behavior may be a byproduct of improvements in mental health that are associated with better health behaviors [21-23]. Alternatively, mood check-ins may increase the perceived benefit of meditation, and thus increase performance of the behavior. Although testing the mechanisms of mood check-ins on meditation behavior was beyond the scope of this paper, this is an important area that warrants additional research. Future studies may benefit from deeper investigation into individuals' experiences related to mood check-ins, such as the potential effects of positive and negative emotions, expectancies about mood regulation [24], self-efficacy, and use of self-feedback to alter thoughts, feelings, and behavior [25,26].

## Limitations

Although this is one of the first studies to assess the effects of mood check-in on participation in meditation, there are important limitations to be noted. First, we did not have any information about subscribers in the current sample beyond their subscription data and app usage patterns. Thus, we did not have information on potentially important covariates that may relate to app usage and meditation behavior and that may modify the effect of mood check-ins on meditation engagement. Future studies are encouraged to collect information on demographics (ie, gender, age, income, education, etc), mental health, and other theoretically important subscriber characteristics to assess the possible confounding impacts of these variables on the relationship between mood check-ins and meditation engagement. Second, we did not assess how users engaged with other behavioral strategies already embedded within the Calm app (ie, tracking time spent meditating, reminders, sharing meditation stats); therefore, it is unknown if mood check-ins operate alone or in tandem with these other features when promoting greater meditation behaviors. Future studies should evaluate the interaction of these behavioral strategies with mood check-ins and their combined effect on participation in app-based meditation. Also, this study analyzed changes in app usage during the first 9 months of users' subscriptions with Calm, but these results may not be generalizable to potential changes in meditation behavior that would occur at other points in a user's subscription. Finally, the cumulative effect of 8 weeks of mood check-ins was an increase of approximately 1

meditation session or 5 minutes of meditation. These findings were statistically significant (both  $P < .001$ ), but it is unclear whether small increases in meditation over an extended period of time constitute meaningful changes associated with clinically significant health benefits. Given the dearth of literature related to meditation dosage, future research is warranted to determine the frequency and duration of meditation optimally needed to convey benefits to participants and whether effects of this magnitude relate to future behavior and health outcomes.

## Conclusions

Using mood check-ins increased meditation participation (in sessions or minutes) and the likelihood of meditation in general, having the greatest benefit for inactive meditators. Mobile apps aiming to encourage better health behaviors should incorporate evidence-based strategies such as mood check-ins to sustain behavior or increase meditation participation, but more research is warranted to examine mechanisms and confirm the findings presented herein. Mobile app developers may also utilize this information to better engage inactive subscribers in an effort to sustain their consumer base.

## Conflicts of Interest

JH conducts investigator-initiated research that utilizes the Calm app, but Calm does not financially support her research. JH is paid for her consultation on an as-needed basis; however, her role is to ensure the quality of Calm's science and she has no specific obligations to the company. JH receives no financial incentives (eg, stocks) related to the growth or success of the company. MP is a paid employee of Behavioral Research and Analytics, LLC.

## Multimedia Appendix 1

Results and figures from supplementary analyses.

[DOCX File, 16603 KB - [mhealth\\_v9i4e27106\\_app1.docx](#) ]

## References

1. 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](#)]
2. Headspace and Snap Inc. team up for world mental health day, launch new headspace mini content to support Snapchatters' rising stress levels. *Business Wire*. 2020 Oct 09. URL: <https://www.businesswire.com/news/home/20201009005034/en/Headspace-and-Snap-Inc.-Team-Up-for-World-Mental-Health-Day-Launch-New-Headspace-Mini-Content-To-Support-Snapchatters?-Rising-Stress-Levels> [accessed 2020-12-14]
3. Tan G. Calm meditation app hits \$2 billion value with lightspeed, goldman backing. *Bloomberg*. 2020 Dec 08. URL: <https://www.bloomberg.com/news/articles/2020-12-08/calm-hits-2-billion-value-with-lightspeed-goldman-tpg-backing> [accessed 2020-12-15]
4. Gál É, Ștefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *J Affect Disord* 2021 Jan 15;279:131-142. [doi: [10.1016/j.jad.2020.09.134](#)] [Medline: [33049431](#)]
5. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J Consult Clin Psychol* 2020 Jan;88(1):1-13. [doi: [10.1037/ccp0000459](#)] [Medline: [31697093](#)]
6. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: [10.2196/14273](#)] [Medline: [31237569](#)]
7. Huberty J, Eckert R, Larkey L, Kurka J, Rodríguez De Jesús SA, Yoo W, et al. Smartphone-based meditation for myeloproliferative neoplasm patients: feasibility study to inform future trials. *JMIR Form Res* 2019 Apr 29;3(2):e12662 [FREE Full text] [doi: [10.2196/12662](#)] [Medline: [31033443](#)]
8. Mobile apps: what's a good retention rate? Upland. URL: <https://uplandsoftware.com/localytics/resources/blog/mobile-apps-whats-a-good-retention-rate/> [accessed 2021-04-16]
9. 21% of users abandon an app after one use. Upland. URL: <https://uplandsoftware.com/localytics/resources/blog/21-percent-of-users-abandon-apps-after-one-use/> [accessed 2021-04-16]
10. Lin Z, Althoff T, Leskovec J. I'll be back: on the multiple lives of users of a mobile activity tracking application. *Proc Int World Wide Web Conf* 2018 Apr;2018:1501-1511 [FREE Full text] [doi: [10.1145/3178876.3186062](#)] [Medline: [29780978](#)]
11. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *Mhealth* 2018 Mar;4:6-6 [FREE Full text] [doi: [10.21037/mhealth.2018.03.02](#)] [Medline: [29682510](#)]
12. Lee K, Kwon H, Lee B, Lee G, Lee JH, Park YR, et al. Effect of self-monitoring on long-term patient engagement with mobile health applications. *PLoS One* 2018 Jul 26;13(7):e0201166 [FREE Full text] [doi: [10.1371/journal.pone.0201166](#)] [Medline: [30048546](#)]

13. Bandura A. Social cognitive theory: an agentic perspective. *Annu Rev Psychol* 2001;52:1-26. [doi: [10.1146/annurev.psych.52.1.1](https://doi.org/10.1146/annurev.psych.52.1.1)] [Medline: [11148297](https://pubmed.ncbi.nlm.nih.gov/11148297/)]
14. Bandura A. Social cognitive theory of self-regulation. *Organ Behav Hum Decis Process* 1991 Dec;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)]
15. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *J Behav Med* 2008 Feb;31(1):23-33. [doi: [10.1007/s10865-007-9130-7](https://doi.org/10.1007/s10865-007-9130-7)] [Medline: [17899351](https://pubmed.ncbi.nlm.nih.gov/17899351/)]
16. Frederick S, Loewenstein G, O'donoghue T. Time discounting and time preference: a critical review. *J Econ Lit* 2002 Jun;40(2):351-401. [doi: [10.1257/002205102320161311](https://doi.org/10.1257/002205102320161311)]
17. Kabat-Zinn J. Mindfulness-based stress reduction (MBSR). *Constructivism Hum Sci* 2003;8(2):73.
18. Chiesa A. Vipassana meditation: systematic review of current evidence. *J Altern Complement Med* 2010 Jan;16(1):37-46. [doi: [10.1089/acm.2009.0362](https://doi.org/10.1089/acm.2009.0362)] [Medline: [20055558](https://pubmed.ncbi.nlm.nih.gov/20055558/)]
19. Caldeira Y, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. *AMIA Annu Symp Proc* 2017;2017:495-504 [FREE Full text] [Medline: [29854114](https://pubmed.ncbi.nlm.nih.gov/29854114/)]
20. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath K, editors. *Health Behavior and Health Education: Theory, Research, and Practice*. San Francisco, CA: Jossey-Bass; 2008:97-121.
21. Kushlev K, Drummond DM, Diener E. Subjective well-being and health behaviors in 2.5 million Americans. *Appl Psychol Health Well Being* 2020 Mar 03;12(1):166-187. [doi: [10.1111/aphw.12178](https://doi.org/10.1111/aphw.12178)] [Medline: [31482675](https://pubmed.ncbi.nlm.nih.gov/31482675/)]
22. Walsh JL, Senn TE, Carey MP. Longitudinal associations between health behaviors and mental health in low-income adults. *Transl Behav Med* 2013 Mar 5;3(1):104-113 [FREE Full text] [doi: [10.1007/s13142-012-0189-5](https://doi.org/10.1007/s13142-012-0189-5)] [Medline: [23997836](https://pubmed.ncbi.nlm.nih.gov/23997836/)]
23. Cohen S, Schwartz JE, Bromet EJ, Parkinson DK. Mental health, stress, and poor health behaviors in two community samples. *Prev Med* 1991 Mar;20(2):306-315. [doi: [10.1016/0091-7435\(91\)90029-4](https://doi.org/10.1016/0091-7435(91)90029-4)]
24. Jimenez SS, Niles BL, Park CL. A mindfulness model of affect regulation and depressive symptoms: positive emotions, mood regulation expectancies, and self-acceptance as regulatory mechanisms. *Pers Individ Dif* 2010 Oct;49(6):645-650. [doi: [10.1016/j.paid.2010.05.041](https://doi.org/10.1016/j.paid.2010.05.041)]
25. Randal C, Pratt D, Bucci S. Mindfulness and self-esteem: a systematic review. *Mindfulness* 2015 Apr 9;6(6):1366-1378. [doi: [10.1007/s12671-015-0407-6](https://doi.org/10.1007/s12671-015-0407-6)]
26. Vago DR, Silbersweig DA. Self-awareness, self-regulation, and self-transcendence (S-ART): a framework for understanding the neurobiological mechanisms of mindfulness. *Front Hum Neurosci* 2012;6:296-296 [FREE Full text] [doi: [10.3389/fnhum.2012.00296](https://doi.org/10.3389/fnhum.2012.00296)] [Medline: [23112770](https://pubmed.ncbi.nlm.nih.gov/23112770/)]

*Edited by L Buis; submitted 12.01.21; peer-reviewed by MS Liew, I Beintner; comments to author 21.01.21; revised version received 23.02.21; accepted 02.04.21; published 23.04.21.*

*Please cite as:*

Huberty J, Green J, Puzia M, Stecher C

*Evaluation of Mood Check-in Feature for Participation in Meditation Mobile App Users: Retrospective Longitudinal Analysis*

*JMIR Mhealth Uhealth* 2021;9(4):e27106

URL: <https://mhealth.jmir.org/2021/4/e27106>

doi: [10.2196/27106](https://doi.org/10.2196/27106)

PMID: [33890865](https://pubmed.ncbi.nlm.nih.gov/33890865/)

©Jennifer Huberty, Jeni Green, Megan Puzia, Chad Stecher. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 23.04.2021. 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.

Original Paper

# An Implementation Strategy to Expand Mobile Health Use in HIV Care Settings: Rapid Evaluation Study Using the Consolidated Framework for Implementation Research

Wendy F Cohn<sup>1</sup>, PhD; Chelsea E Canan<sup>2</sup>, MPH, PhD; Sarah Knight<sup>2</sup>, MPH, CHES; Ava Lena Waldman<sup>2</sup>, MHS, CHES, CCRP; Rebecca Dillingham<sup>2</sup>, MD, MPH; Karen Ingersoll<sup>3</sup>, PhD; Julie Schexnayder<sup>2</sup>, MPH, DNP; Tabor E Flickinger<sup>4</sup>, MD, MPH

<sup>1</sup>Department of Public Health Sciences, University of Virginia Cancer Center, University of Virginia, Charlottesville, VA, United States

<sup>2</sup>Division of Infectious Diseases, Department of Medicine, University of Virginia, Charlottesville, VA, United States

<sup>3</sup>Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA, United States

<sup>4</sup>Division of General, Geriatric, Palliative and Hospital Medicine, Department of Medicine, University of Virginia, Charlottesville, VA, United States

**Corresponding Author:**

Wendy F Cohn, PhD

Department of Public Health Sciences

University of Virginia Cancer Center

University of Virginia

560 Ray C Hunt Dr

PO Box 800765

Charlottesville, VA, 22908

United States

Phone: 1 434 964 7811

Email: [wfc2r@virginia.edu](mailto:wfc2r@virginia.edu)

## Abstract

**Background:** Mobile health (mHealth) apps can provide support to people living with a chronic disease by offering resources for communication, self-management, and social support. PositiveLinks (PL) is a clinic-deployed mHealth app designed to improve the health of people with HIV. In a pilot study, PL users experienced considerable improvements in care engagement and viral load suppression. To promote its expansion to other HIV clinics, we developed an implementation strategy consisting of training resources and on-demand program support.

**Objective:** The objective of our study was to conduct an interim analysis of the barriers and facilitators to PL implementation at early adopting sites to guide optimization of our implementation strategy.

**Methods:** Semistructured interviews with stakeholders at PL expansion sites were conducted. Analysis of interviews identified facilitators and barriers that were mapped to 22 constructs of the Consolidated Framework for Implementation Research (CFIR). The purpose of the analysis was to identify the facilitators and barriers to PL implementation in order to adapt the PL implementation strategy. Four Ryan White HIV clinics were included. Interviews were conducted with one health care provider, two clinic managers, and five individuals who coordinated site PL activities.

**Results:** Ten common facilitators and eight common barriers were identified. Facilitators to PL implementation included PL's fit with patient and clinic needs, PL training resources, and sites' early engagement with their information technology personnel. Most barriers were specific to mHealth, including access to Wi-Fi networks, maintaining patient smartphone access, patient privacy concerns, and lack of clarity on how to obtain approvals for mHealth use.

**Conclusions:** The CFIR is a useful framework for evaluating mHealth interventions. Although PL training resources were viewed favorably, we identified important barriers to PL implementation in a sample of Ryan White clinics. This enabled our team to expand guidance on identifying information technology stakeholders and procuring and managing mobile resources. Ongoing evaluation results continue to inform improvements to the PL implementation strategy, facilitating PL access for future expansion sites.

(*JMIR Mhealth Uhealth* 2021;9(4):e19163) doi:[10.2196/19163](https://doi.org/10.2196/19163)



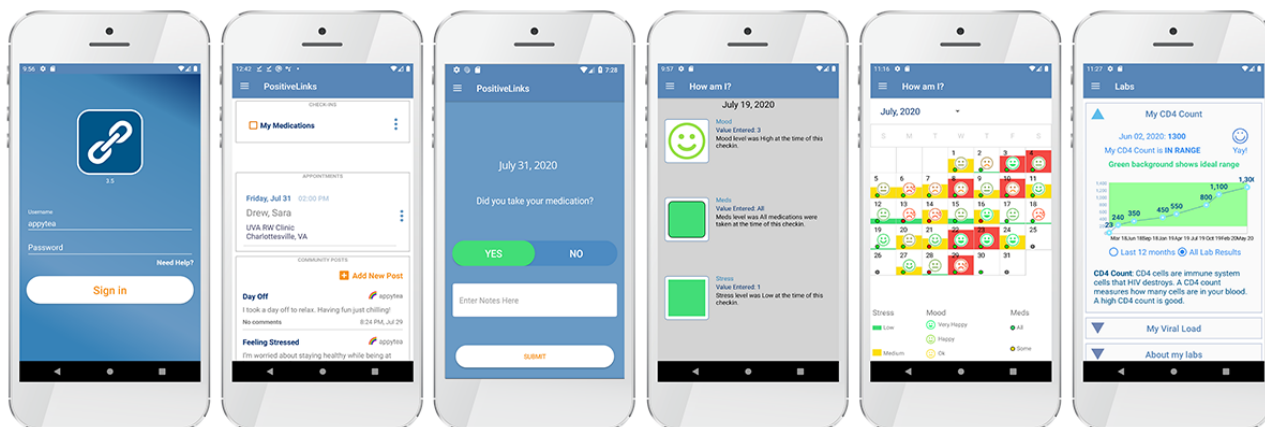
## KEYWORDS

mHealth; smartphone; mobile health; implementation strategy; implementation science; Consolidated Framework for Implementation Research; HIV care engagement; viral suppression

## Introduction

People living with HIV achieve positive health outcomes more effectively when they establish and maintain primary HIV care and have high adherence to their antiretroviral therapy [1-3]. Mobile health (mHealth) apps can support chronic disease self-management in this population by providing a platform for communication, self-monitoring, and social support. Improved outcomes have been reported for mHealth users with chronic diseases including asthma, diabetes, and HIV [4-8]. Because of their potential for improving health outcomes, health systems are beginning to leverage mHealth interventions to engage people with HIV in self-management behaviors [9]. Accordingly, there is an unmet need for implementation strategies to support bidirectional mHealth use by people living with HIV and their primary care teams.

**Figure 1.** Positive Links screen shots: patient app.



PL is administered by the HIV primary care clinic. At the University of Virginia, PL is coordinated by a dedicated team member who enrolls patients and HIV care team members in the program, issues smartphones and payments for smartphone services to patients, provides patient training on use of the app's features, manages app content, and assists with phone and app troubleshooting. Selected clinical laboratory results and other medical information are directly imported from the electronic medical record to PL for patient viewing within the app. Referrals to the program are made by clinical and nonclinical members of the HIV care team. HIV care team members access PL as needed to exchange messages with patients in a password-protected environment and, if desired by team members, to monitor their patients' self-reported medication adherence. Patients are asked to complete daily check-ins on their HIV medication use, moods, and stress, and have unrestricted access to all app features.

In 2017, PL was made available to other HIV service providers. The primary targets for PL expansion were HIV clinics that were funded under the Health Resources and Services Administration's Ryan White HIV/AIDS Program. The objectives of this study were to describe an interim rapid evaluation of PL implementation determinants at early adopting

*PositiveLinks* (PL) is a clinic-based mHealth intervention that was designed in partnership with people living with HIV to increase engagement in care and improve clinical outcomes [10]. PL consists of a smartphone app for patients that includes features, such as appointment reminders, a virtual community board, and daily queries of mood, stress, and medication adherence, with graphical feedback of behavioral patterns (Figure 1) paired with a suite of tools for clinics and providers. Following app development, PL was piloted at the University of Virginia's Ryan White Clinic in a single arm prospective study. We found considerable improvements in engagement in care and in key clinical laboratory markers (CD4 count and HIV viral suppression), measures of immune system recovery, and cessation of HIV viral replication. The app was subsequently adopted as usual care at this site [11].

Ryan White clinics using the Consolidated Framework for Implementation Research (CFIR) and to describe the process for refining our PL implementation support program based on the study findings.

## Methods

### Setting

The University of Virginia Institutional Review Board approved this study. The setting for this study was four Ryan White clinics that pursued PL implementation in 2018. Ryan White clinics provide HIV primary medical care, medications, and essential support services for people living with HIV who are low income, uninsured, and underserved [12]. The clinics implementing PL varied in their location and organizational structure. They included three health system-affiliated Ryan White clinics in Virginia and one in Texas. All four clinics were located in cities, with the two largest clinics serving primarily an urban population and the two remaining clinics reaching a catchment area that included both urban and rural regions. At the time of this evaluation, of the four sites, one ceased participation in PL before enrolling patients due to difficulties with leadership buy-in, one was unable to obtain information security approvals

for PL use within the parent health system, and two progressed to the implementation stage and began enrolling patients.

All sites received PL implementation support from the University of Virginia PL implementation team that incorporated evidence-based interventions from an implementation research taxonomy [13]. Implementation support included a comprehensive training package consisting of a training manual, onsite PL training for HIV care team members, a PL learning management system accessible directly from PL's clinic/provider tool suite, and ongoing and on-demand program support from an experienced PL coordinator.

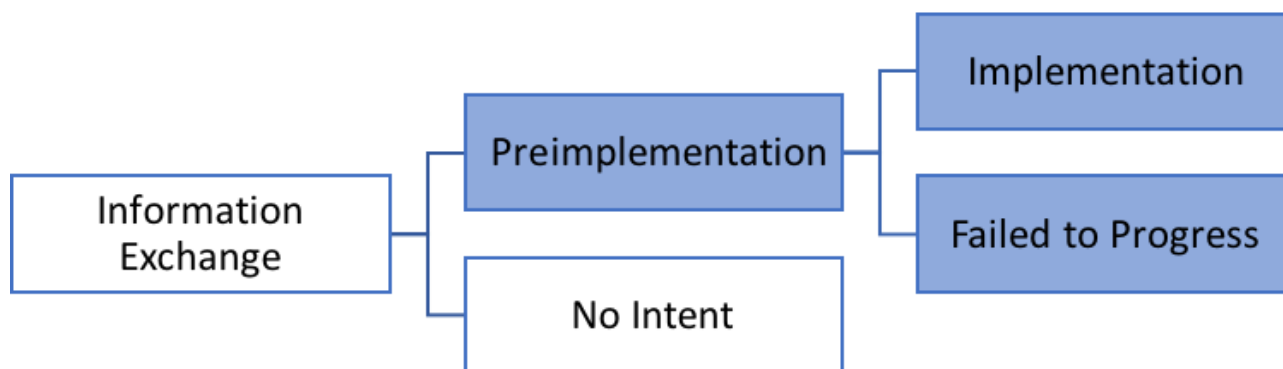
We anticipated that successful implementation of PL would require processes and infrastructure on multiple levels. At the individual level, PL members (individuals living with HIV) must download and install the app and interact with the app for self-monitoring. At the interpersonal level, PL members communicate with the implementation site-specific PL coordinator and their care team. PL coordinators at the sites also interact with the University of Virginia PL implementation team for training and support. PL workflow must be integrated into the clinic site and processes developed for allocation of phone resources, member enrollment, and support and engagement for members and providers. At the organizational

level, the PL program requires leadership approval and information technology (IT) infrastructure to support it.

### Recruitment and Study Participants

Purposive sampling was used to recruit employees from the Ryan White clinics that expressed intent to adopt the PL program. Intent to adopt PL was ascertained from direct email inquiries, personal contact with the investigators, or requests for information on the PL website. The PL implementation coordinator at the University of Virginia tracked these clinics in their progression toward PL implementation. Figure 2 displays the stages of the implementation pathways from *information exchange* when potential sites contact PL to learn about the program to *preimplementation* after the decision is made to implement PL and then to either *implementation* or *failed to progress*. Sites were included in the evaluation when the clinic made the decision to adopt PL (preimplementation). The evaluation team was notified to initiate interview recruitment activities and received contact information for the individual confirming intent to implement PL. This individual was contacted by phone or email and asked to identify primary stakeholders involved with the PL implementation process at the site. The individuals were then approached to participate in the study via email.

**Figure 2.** Stages of implementation. Interviews occurred during the stages shaded in blue.



We specifically targeted clinic managers, PL coordinators, and PL providers for interviews. PL coordinators were the point people for PL at each site; they were responsible for enrollment and support of PL members. We defined PL providers as physicians, nurses, psychologists, social workers, case managers, community health workers, or any other staff members who were end-users of the PL platform.

### Implementation Framework Selection

An implementation science determinants framework informed the interview guide development and analysis, and it allowed for consideration of the multilevel factors that may contribute to implementation success [14,15]. The CFIR [16] captures a broad range of constructs that fit the goals of this evaluation and provides interview guides and codebook templates that facilitate the application of this framework to evaluation projects [17]. At the time of our evaluation, the CFIR was being used

widely in health-related implementation research [18-27] but had a smaller presence in the field of mHealth [28-31].

The CFIR consists of the following five domains: *intervention characteristics*, *outer setting*, *inner setting*, *characteristics of individuals*, and *implementation process*. Across the primary domains are 39 specific constructs corresponding to factors of successful implementation [16,17]. The CFIR is designed to be flexible in its use, with users able to determine which constructs are most relevant to their project's implementation.

Two sources guided the selection of specific CFIR constructs for inclusion in the evaluation. The University of Virginia PL implementation team reviewed and identified all five domains and six priority constructs (innovation source, evidence strength and quality, patient needs and resources, self-efficacy, and engaging opinion leaders) as most relevant to the PL implementation based on their experiences with early adopting sites. Additionally, a 2016 systematic review summarized factors

that influence the implementation of all types of eHealth interventions [31]. We used findings from the systematic review to select CFIR constructs most relevant to mHealth interventions. Unlike mHealth interventions described in the report by Ross et al [31], cost was not anticipated to be a barrier to initial PL implementation. All four Ryan White clinics had received grant funding to support their PL programs. In total, 22 CFIR constructs were included in the analysis. The chosen constructs with their PL-specific operational definitions are included in Table 1. We also included an additional code for *technology* because the mHealth aspect of the intervention was particularly salient to our implementation, and we wanted to be sure that there was a code to quickly capture all data related to this theme.

### Interview Guide Development

We developed three interview guides to permit tailoring of questions to clinic managers, PL coordinators, and PL providers. PL coordinators were asked questions about the PL program at two time points (before implementation and during implementation) using the same interview guide. PL providers were asked questions about the PL program during implementation only (ie, after the clinic began enrolling patients in the PL program). Interview guides incorporated questions associated with 22 CFIR constructs (Table 1). For each construct, we began by adapting questions suggested by the CFIR developers [17] and added questions as necessary for our particular project needs.

**Table 1.** Included Consolidated Framework for Implementation Research constructs and their operationalization.

Construct	Before implementation	During implementation	
	Coordinator	Coordinator	Provider
<b>Innovation characteristics</b>			
Innovation source <i>“Who developed PL<sup>a</sup>? Why is PL being implemented in your clinic?”</i>	No	No	Yes
Evidence strength and quality <i>“What evidence are you aware of that shows whether PL will work in your clinic?”</i>	No	No	Yes
Adaptability <i>“What changes will you need to make to PL so it works in your clinic?”</i>	Yes	No	No
Complexity <i>“How complicated is PL?”</i>	No	No	Yes
Design quality and packaging <i>“What is your perception of supporting materials, packaging and bundling of PL?”</i>	No	Yes	Yes
<b>Outer setting</b>			
Patient needs and resources <i>“How well does PL meet the needs of your patients? How do patients respond to PL?”</i>	No	Yes	Yes
External policy and incentives <i>“What local, state, national policies or guidelines influenced your decision to implement PL?”</i>	Yes	No	No
<b>Inner setting</b>			
Compatibility <i>“How does PL fit with the values and norms in your clinic? How does PL fit into clinic processes and workflow?”</i>	Yes	No	Yes
Leadership engagement <i>“What kind of support for PL have you seen from leaders in your clinic?”</i>	Yes	Yes	No
Available resources <i>“What resources do you need to implement PL? Do you have those resources?”</i>	Yes	No	Yes
Access to knowledge and information <i>“What kind of training is planned for you and your colleagues?”</i>	Yes	No	Yes
<b>Characteristics of individuals</b>			
Knowledge and beliefs about the intervention <i>“Do you think PL will be effective? How do you feel about your plan to implement PL?”</i>	Yes	No	Yes
Self-efficacy <i>“How confident do you feel about implementing PL?”</i>	Yes	No	Yes
Other personal attributes <i>“Tell me about yourself and your role with PL”</i>	Yes	No	Yes
<b>Implementation process</b>			
Planning <i>“To what extent is there a plan in place to implement PL? Who is involved? What role has your plan played in implementation?”</i>	Yes	Yes	No

Construct	Before implementation	During implementation	
	Coordinator	Coordinator	Provider
Engaging opinion leaders “Who are key individuals to get on board? What are they saying about PL?”	Yes	Yes	No
Engaging formally appointed internal implementation leaders “Who will lead PL implementation? How did your clinic get involved in PL?”	Yes	No	No
Engaging champions “Are there people who go above and beyond what might be expected?”	Yes	No	No
Engaging key stakeholders “What steps are taken to encourage participation in PL?”	No	No	Yes
Engaging innovation participants “How do you communicate PL to patients?”	No	No	Yes
Executing “Has PL been implemented according to plan?”	No	Yes	Yes
Reflecting and evaluating “What kind of information do you collect as you implement PL? How do you assess progress toward your goals?”	Yes	Yes	No

<sup>a</sup>PL: PositiveLinks.

## Interviews

Recruitment began at all clinics on February 15, 2018. Interviews were completed between March 1, 2018, and July 10, 2019. All participants verbally provided their informed consent to participate and were allowed to discontinue participation at any time after giving consent. Interviews were conducted over the phone and were audio recorded and professionally transcribed. Interview lengths differed based on the guide used, ranging from 30 to 60 minutes. Interviewers were trained members of the evaluation team, and they were not involved with PL implementation at the clinics. This was done to promote candid and honest responses from the interview participants when asked to describe their experiences. Interviews were stored as audio files on a secure drive labeled by study ID number. Individual interviews were not discussed with the entire team to preserve confidentiality.

Eight interviews were completed, including three interviews that occurred during the preimplementation stage and five interviews that occurred after the clinics began enrolling patients in their PL programs. Six of the interviews were contributed by two clinics that successfully implemented PL during the study period. The remaining clinics were unable to implement PL as of June 30, 2019. Each of these clinics contributed a single interview. One of the clinics experienced difficulties in garnering leadership buy-in for PL use. The other clinic was unable to obtain information security approvals for PL use within the parent health system.

Because only two sites progressed to enrolling patients in PL, PL providers were only eligible to participate in interviews at

two sites. Of 13 providers we attempted to recruit via email, two declined interviews, one was lost to follow-up, nine did not respond, and one successfully completed the interview.

## Analysis

An analytical template was used by trained personnel (WC, CC, JS, and TEF) using the CFIR codebook [17], which provided operational definitions for CFIR constructs along with example inclusion and exclusion text. We restricted the codebook to those constructs selected for inclusion as described above, with the addition of the *technology* code as previously described. Transcripts were independently coded by two investigators who then worked together to achieve consensus on the coded content. Coding and analyses were performed using Dedoose Version 8.2.14 [32].

After reaching consensus on the coded content, the two coders independently summarized the barriers and facilitators that emerged from each interview. These summaries also underwent consensus discussions. The summaries were used to generate a master list of PL implementation determinants. Each determinant was listed with its corresponding CFIR construct-specific examples from the interviews and potential action items to report to the implementation team. Facilitators were evaluated by the team to assess whether there were corresponding actions that would enhance or strengthen the facilitator and enable future clinics to benefit from explicit recommended actions.

The master list of PL implementation determinants, including facilitators and barriers, was updated and shared with the implementation team as each determinant summary was

completed. This rapid evaluation process enabled iterative changes to the implementation model prior to study completion. The median time from interview completion to determinant summary completion was 131 days and ranged from 36 to 228 days. All information presented to the implementation team was delivered in a deidentified and aggregated manner.

## Results

### Facilitators

Based on our analyses, nine CFIR constructs were associated with 10 facilitators. A summary of facilitators and recommended action items is shown in [Table 2](#). *PL compatibility* was an important facilitator to its implementation. Components of fit included alignment between (1) PL and clinic needs, and (2) PL and clinic goals and values. The intersection between *patient needs and resources* and *compatibility* constructs was notable. For example, the respondents mentioned that accessing the clinic was a challenge for some of their patients. Having a phone and connectivity to the clinic through PL helped to address this barrier. The following two action items arose from findings related to PL compatibility: (1) remind sites that it is important to budget for cell phones for their most at risk patients and (2)

help sites identify their own needs and articulate how PL addresses those needs during preimplementation.

Engaging key stakeholders was also identified as an important facilitator of PL implementation. First, stakeholder engagement with PL implementation activities contributed to leadership and end-user excitement for the program. Second, stakeholder input during implementation planning prepared teams to integrate PL into their existing workflows. Encouraging sites to include clinic staff in implementation planning was identified as an action item.

The *planning* and *engaging innovation participants* constructs identified useful strategies for boosting engagement by PL providers and users. Action items included proactive planning for these engagement strategies and timing marketing of PL features that are dependent on group participation. The remaining CFIR constructs (and their associated facilitators) highlighted aspects of the implementation process that were going well and did not require additional action. Examples include accessibility and quality of PL training materials, ease of PL use, and PL's ability to be adapted to clinic needs. For these facilitators, the evaluation team recommended that the implementation team continue current practices.

**Table 2.** Facilitators to PositiveLinks implementation.

CFIR <sup>a</sup> domain and construct: facilitator	Example	Action item
<b>Outer setting</b>		
Needs and resources of those served: Perceived match between needs of clients and PL <sup>b</sup> features	<ul style="list-style-type: none"> <li>• PL perceived as meeting patient needs for engagement, communication, support, medication adherence, appointments, and lab tracking</li> <li>• Staff believe PL can help patients who have difficulty getting to the clinic</li> <li>• Phones help patients stay in touch with the clinic and family/friends</li> </ul>	<ul style="list-style-type: none"> <li>• Remind sites of the importance of budgeting for cell phones for most at risk patients</li> <li>• Help sites identify their own needs and then discuss how PL can address these needs</li> </ul>
<b>Inner setting</b>		
Compatibility: PL meets the needs of the clinic/staff	<ul style="list-style-type: none"> <li>• PL is supporting what staff are already doing to engage patients in care</li> <li>• Incorporating PL into clinic operations and quality management plans</li> <li>• PL perceived as helping to overcome communication-related gaps in engagement</li> </ul>	<ul style="list-style-type: none"> <li>• Emphasize that PL might make it easier for staff to do what they are already doing</li> <li>• Remind sites that this tool was developed to meet the needs identified by clinicians</li> </ul>
Compatibility: PL alignment with clinic goals and values	<ul style="list-style-type: none"> <li>• Good alignment between goals of the clinic and PL: connecting to clients, medication adherence, and patient-centered focus</li> </ul>	<ul style="list-style-type: none"> <li>• Ask clinics to identify their values and goals, for example, setting targets for retention-in-care or viral suppression rates that may be improved by PL use.</li> </ul>
Access to knowledge and information: Quality of PL training materials	<ul style="list-style-type: none"> <li>• Positive impression of training, materials, and support for both learning the program and navigating through the approval process</li> <li>• Plan for training is well developed, occurs at an appropriate time, and is delivered to the right staff</li> </ul>	<ul style="list-style-type: none"> <li>• No action indicated</li> </ul>
<b>Innovation characteristics</b>		
Complexity: Ease of PL use	<ul style="list-style-type: none"> <li>• Simplicity and user friendliness of the patient-facing app</li> <li>• Web portal viewed as simple and easy; made it easier for staff to use PL</li> <li>• Web portal includes metrics desired by the clinic for the tracking program</li> </ul>	<ul style="list-style-type: none"> <li>• No action indicated</li> </ul>
Adaptability: Ability to adapt PL to unique clinic workflows	<ul style="list-style-type: none"> <li>• Ability to tailor PL, such as who receives PL messages</li> <li>• Ability to adapt the web portal to show desired information</li> </ul>	<ul style="list-style-type: none"> <li>• No action indicated</li> </ul>
<b>Implementation process</b>		
Engaging key stakeholders: Function and roles of the clinic team	<ul style="list-style-type: none"> <li>• Leadership at the clinic (CEO and clinic supervisor) is committed to the program</li> <li>• Teamwork within the site to identify clients likely to benefit from PL and prioritize their enrollment</li> <li>• Evolution of roles over time, that is, the supervisor has more responsibility during the approval phase and then responsibility transfers to coordinators</li> </ul>	<ul style="list-style-type: none"> <li>• Consider creating an opportunity for coordinators at different sites to interact with each other and share their experiences in order to build engagement as a community of practice</li> </ul>

CFIR <sup>a</sup> domain and construct: facilitator	Example	Action item
Planning: Planning	<ul style="list-style-type: none"> <li>• Teams feel better prepared for PL implementation when having a plan of who to enroll first, who will manage PL, and how PL will fit into their workflow</li> <li>• Proactive engagement with information technology security, anticipating the need for key approvals and proactively seeking them</li> <li>• Soft launch with trial run, including mock patients and messages, to get clinic staff engaged and comfortable</li> <li>• Create plans with milestones and timelines</li> </ul>	<ul style="list-style-type: none"> <li>• Plan for clinician and other clinic staff engagement by asking for their input for program improvement</li> </ul>
Engaging innovation participants: Initial success of rollout	<ul style="list-style-type: none"> <li>• Early success during implementation (clinicians buy-in; clients loving it)</li> </ul>	<ul style="list-style-type: none"> <li>• Emphasize individual-level features first (check-ins, resources) and phase in the community board when there are enough participants to make it engaging</li> </ul>
<b>Characteristics of individuals</b>		
Knowledge and beliefs about the innovation: Perceptions of PL	<ul style="list-style-type: none"> <li>• Positive attitudes toward PL and its implementation by the PL coordinator, providers, and other staff</li> </ul>	<ul style="list-style-type: none"> <li>• No action indicated</li> </ul>

<sup>a</sup>CFIR: Consolidated Framework for Implementation Research.

<sup>b</sup>PL: PositiveLinks.

## Barriers

Six CFIR constructs were associated with eight common barriers. A summary of those barriers and their recommended action items are shown in Table 3. Barriers associated with PL *compatibility* were common (n=3) and related to mHealth technology either directly or indirectly. Specific examples expressed by the respondents included not being able to get Wi-Fi in the clinic to download PL on their phones and the release of laboratory results on PL prior to review by the medical provider, which was in opposition to some clinic's usual workflows.

The remaining barriers were associated with different CFIR constructs. *Policy and incentives* represented a single barrier related to privacy concerns, both from clients and administrators

concerned about regulatory compliance. Issues in implementation *planning* arose in relation to obtaining institutional approvals for PL use, in part because of unclear internal review processes. Barriers associated with *available resources* focused on phone availability and staffing effort to manage the PL program. PL's *adaptability* by clinics created potential barriers to uptake, with suggestions for allowing tailoring of the PL platform appearance to appeal to unique clinic populations. To address these barriers, the team developed the following action items: (1) create a document that helps sites anticipate potential IT challenges with tips on how to address in advance and (2) provide examples from other sites, including stories or case studies related to how sites addressed common problems and how long their implementation process took.



**Table 3.** Barriers to PositiveLinks implementation.

CFIR <sup>a</sup> domain and construct: barrier	Example	Action item
<b>Outer setting</b>		
External policy and incentives: Privacy	<ul style="list-style-type: none"> <li>• Privacy concerns from information privacy officers</li> <li>• Clients are concerned about privacy issues</li> </ul>	<ul style="list-style-type: none"> <li>• Create a document with suggestions to help sites prepare for anticipated challenges. Include tips such as figuring out who key decision makers are, what permissions are needed, and identifying all the people the team will likely need on board (eg, privacy, security, and clinical)</li> </ul>
<b>Inner setting</b>		
Compatibility: Information technology	<ul style="list-style-type: none"> <li>• Wi-Fi access at enrollment locations</li> <li>• Phone related (permission to trust app, troubleshooting phone technology)</li> <li>• PL not integrated with EMRb</li> </ul>	<ul style="list-style-type: none"> <li>• PL<sup>c</sup> prioritizes EMR integration</li> </ul>
Compatibility: Mismatch of goals/priorities	<ul style="list-style-type: none"> <li>• Clinicians focused on benefits to patients; may not be aware of PL goals set by an external decision maker</li> <li>• Mismatch between desire of the clinic director to implement PL and the buy-in from staff carrying out the implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Develop new strategies for communicating among site stakeholders about goals and priorities</li> </ul>
Compatibility: Clinic workflow/structure	<ul style="list-style-type: none"> <li>• Concern about patients seeing their lab results in PL before their appointment</li> <li>• Competing priorities for clinic staff</li> </ul>	<ul style="list-style-type: none"> <li>• Allow sites to tailor the lab feed to meet their own needs; consider only releasing lab values after provider review</li> <li>• Emphasize that clinics consider multiple responsibilities of staff and discuss ways to fit PL into existing workflow based on their unique processes</li> </ul>
Available resources: Resources	<ul style="list-style-type: none"> <li>• Acquiring cell phones and coordinating cell phone payments</li> <li>• Limited resources to handle enrollments, manual entry of lab results, and appointments in PL</li> <li>• Staffing numbers and capacity to successfully enact an mHealth intervention</li> </ul>	<ul style="list-style-type: none"> <li>• Continue sharing the reference document outlining the different phone service providers and how to pay them</li> <li>• Create a learning module specific to cell phone payments; provide a customizable template for sites</li> <li>• Re-emphasize staffing needs, including that needs may change over time</li> </ul>
<b>Implementation process</b>		
Planning: Preimplementation approvals	<ul style="list-style-type: none"> <li>• Unclear how to initiate internal approval processes</li> <li>• Decisions to adopt PL disconnected from PL users</li> <li>• Initial concern over the mechanics and length of time needed to implement PL</li> </ul>	<ul style="list-style-type: none"> <li>• Provide examples from other sites</li> <li>• Consider writing stories/case studies about implementation processes that have succeeded at other sites, including how long it takes to go through each step of the process</li> </ul>

CFIR <sup>a</sup> domain and construct: barrier	Example	Action item
Engaging key stakeholders: Buy-in	<ul style="list-style-type: none"> <li>• Doctors see PL as “another thing to do”</li> <li>• Coordinator unsure of the level of buy-in from all staff including nurses</li> <li>• Coordinators need targeted engagement strategies to keep clients and providers using the app</li> </ul>	<ul style="list-style-type: none"> <li>• Consider providing literature on provider engagement</li> <li>• Remind coordinators that PL can still be beneficial to patients if providers do not engage in the app</li> <li>• Clearly articulate that “providers” can include other staff roles, not only doctors</li> <li>• Consider retraining providers if there is uneven engagement or high turnover</li> <li>• Consider involving the “frontline” staff in early implementation decision making and planning to improve engagement, motivation, and compatibility</li> <li>• Set realistic goals for PL participation; provide guidance on what goals to aim for and how to track goals</li> </ul>
<b>Innovation characteristics</b>		
Adaptability: Adaptability and design of PL	<ul style="list-style-type: none"> <li>• PL is developed externally by a site that is different than the expansion site</li> <li>• Some PL features do not meet the preferences of clients (older patients may have difficulty or lack of interest in a mobile app; younger patients may prefer a more upgraded interface)</li> </ul>	<ul style="list-style-type: none"> <li>• Anticipate needing to adapt PL from one clinic population to another; seek more input from the clinic staff up front about their clients’ needs</li> <li>• Ensure that app updates and upgrades continue on an ongoing basis following feedback from users</li> </ul>

<sup>a</sup>CFIR: Consolidated Framework for Implementation Research.

<sup>b</sup>EMR: emergency medical record.

<sup>c</sup>PL: PositiveLinks.

## Discussion

### Principal Findings

This work demonstrates the feasibility of applying the CFIR to the evaluation of mHealth implementations. Rapid evaluation methods using a determinants framework were deemed ideal for interim analysis, simultaneously providing for rigorous assessment of PL implementation processes, identification of specific barriers and facilitators of implementation, and timely refinement of our implementation support program [14].

One of the goals of this analysis was to determine if CFIR could adequately capture the important factors in mHealth implementation. We found that the constructs within CFIR are sufficient to pick up mHealth-specific considerations that may impact successful implementation. Although most content captured within our functional *technology* code was also coded with a CFIR construct, its inclusion in our final codebook enabled more rapid data reduction and analysis to identify instances of delayed or stalled PL uptake due to the clinic’s or health system’s technology-related barriers. Technology played a role in each of the CFIR domains. The addition of a functional code for *technology* allowed the evaluation team to extract segments of interviews specific to technology aspects of the implementation support program. This enabled us to target our rapid analysis; however, CFIR alone is suitable to cover the concepts arising in mHealth intervention-related interviews.

Preliminary results from the first four PL expansion sites identified *compatibility*, *engaging key stakeholders*, and

*innovation participants and planning* as important CFIR constructs associated with early PL implementation. The primary barriers identified in our early results were related to technology and mHealth. mHealth interventions require support from a broad range of stakeholders, including clinic leadership and administration, clinic staff, security and privacy officers, and IT personnel. One recommendation that emerges from this finding is that sites intending to implement an mHealth intervention should engage their security and technology leaders and staff early in the process to streamline implementation and help to avoid or minimize technology-related barriers, such as multiple levels of review and permissions that can be required by large systems. Further, mHealth interventions are unique in their need for continuous updates following their initial implementation. Unlike more discrete interventions, mHealth interventions are not complete once they are implemented but instead require ongoing technological maintenance and support. We identified a useful application of the CFIR for identifying barriers and facilitators at PL sites while PL implementation processes were underway. This enabled us to provide specific action items for the PL implementation team, and it resulted in iterative refinement of our PL implementation strategy. The evaluation was successful in providing interim CFIR-informed feedback to implementation stakeholders rather than waiting for the end of implementation to assess its success. This model could be useful in other implementations that occur on a rolling or ongoing basis.

The results of our evaluation led to revisions in a detailed implementation manual that is provided to new expansion sites.

This manual includes information about the PL program itself, including a detailed description of all app features and components of the web portal. Revisions to the implementation manual included (1) a description of important milestones in the implementation process including developing a budget and IT flows, (2) information about gathering necessary security and privacy approvals, and (3) phone logistics. This manual is essential to the successful implementation of PL at new sites and is continually updated based on feedback learned from the implementation interviews using the rapid feedback approach taken by the evaluation and implementation teams.

Rapid evaluation was critical to providing timely feedback to the implementation team and expansion sites. Because the implementation team is continuously implementing PL and working to meet the immediate needs of expansion sites, it is important to plan regular communication between the evaluation and implementation teams. The frequency of communication should be based on timing of interviews and the emergence of new sites. Both teams should establish an integrated flow that allows for two-way channels of communication from the implementation team to the evaluators regarding the status of pending sites, as well as from the evaluators to the implementers regarding recommended changes to the implementation process.

The results of this rapid interim evaluation suggest that health systems looking to adopt new mHealth apps to improve patient engagement and outcomes will need to consider adopting a streamlined approach to decision making and IT infrastructure, including security, to have successful mHealth implementation. Our findings are consistent with other recent evaluations of mHealth interventions [33,34] that used CFIR to identify factors influencing implementation, including characteristics of the innovations themselves, as well as the local and institutional contexts in which they are being adopted. Scale-up of technological innovation is a challenging process and can be inhibited by organizational factors external to the intervention.

Future work will include incorporating interviews with patients who represent another important stakeholder group in the

implementation of mHealth. Additional next steps will include incorporating the CFIR valence attribution as the number of sites increase and we become able to identify the positive and negative attributions of each implementation.

### Limitations

One of the main limitations of this evaluation was the lack of provider interviews. Our provider recruitment experience echoes others in the implementation literature, requiring perseverance of our PL evaluation team and considerably longer recruitment times than predicted [35]. Providers represent a key stakeholder group in mHealth chronic disease interventions, and their participation in the evaluation of such interventions is critical. For interventions like PL that are designed to improve connections between users with chronic medical conditions and members of their health care teams, intervention effectiveness is dependent on provider uptake. Engaging providers in evaluations of PL implementation is essential for the recognition of and response to the unique barriers to PL use.

### Conclusion

This study describes the use of the CFIR to guide iterative refinement of an implementation strategy to facilitate dissemination of our mHealth intervention. Our findings highlight the unique characteristics of mHealth interventions and the multilevel factors that must be considered when planning for their implementation in health care settings. The flexibility and comprehensiveness of CFIR appear to be sufficient to capture concepts within the interviews that we conducted and likely would be applicable to the evaluation of other mHealth interventions. Strategies for rapid evaluation may be particularly important in the realm of mHealth, where the field can move quickly. Rapid evaluation methods that are rigorous and responsive to the experiences of early mHealth adopters can better inform best practices for mHealth implementation. Increasing provider feedback will also enable more impactful and utility-focused evaluation.

---

### Acknowledgments

This study was funded by the Virginia Department of Health. The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the paper, and the final contents.

---

### Conflicts of Interest

RD, KI, and ALW report providing consulting services to Warm Health Technology, Inc. The other authors have no conflicts to declare.

---

### References

1. Mugavero MJ, Amico KR, Westfall AO, Crane HM, Zinski A, Willig JH, et al. Early retention in HIV care and viral load suppression: implications for a test and treat approach to HIV prevention. *J Acquir Immune Defic Syndr* 2012 Jan 01;59(1):86-93 [FREE Full text] [doi: [10.1097/QAI.0b013e318236f7d2](https://doi.org/10.1097/QAI.0b013e318236f7d2)] [Medline: [21937921](https://pubmed.ncbi.nlm.nih.gov/21937921/)]
2. Ortego C, Huedo-Medina TB, Llorca J, Sevilla L, Santos P, Rodríguez E, et al. Adherence to highly active antiretroviral therapy (HAART): a meta-analysis. *AIDS Behav* 2011 Oct;15(7):1381-1396. [doi: [10.1007/s10461-011-9942-x](https://doi.org/10.1007/s10461-011-9942-x)] [Medline: [21468660](https://pubmed.ncbi.nlm.nih.gov/21468660/)]

3. Fleishman JA, Yehia BR, Moore RD, Korthuis PT, Gebo KA, Research Network HIV. Establishment, retention, and loss to follow-up in outpatient HIV care. *J Acquir Immune Defic Syndr* 2012 Jul 01;60(3):249-259 [FREE Full text] [doi: [10.1097/QAI.0b013e318258c696](https://doi.org/10.1097/QAI.0b013e318258c696)] [Medline: [22531758](https://pubmed.ncbi.nlm.nih.gov/22531758/)]
4. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. *Cochrane Database Syst Rev* 2012 Dec 12;12:CD007459 [FREE Full text] [doi: [10.1002/14651858.CD007459.pub2](https://doi.org/10.1002/14651858.CD007459.pub2)] [Medline: [23235644](https://pubmed.ncbi.nlm.nih.gov/23235644/)]
5. 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](https://doi.org/10.1371/journal.pmed.1001362)] [Medline: [23349621](https://pubmed.ncbi.nlm.nih.gov/23349621/)]
6. 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):e52 [FREE Full text] [doi: [10.2196/jmir.3951](https://doi.org/10.2196/jmir.3951)] [Medline: [25803266](https://pubmed.ncbi.nlm.nih.gov/25803266/)]
7. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, 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](https://doi.org/10.2196/mhealth.8873)] [Medline: [29343463](https://pubmed.ncbi.nlm.nih.gov/29343463/)]
8. van Uden-Kraan CF, Drossaert CHC, Taal E, Seydel ER, van de Laar MAFJ. Participation in online patient support groups endorses patients' empowerment. *Patient Educ Couns* 2009 Jan;74(1):61-69. [doi: [10.1016/j.pec.2008.07.044](https://doi.org/10.1016/j.pec.2008.07.044)] [Medline: [18778909](https://pubmed.ncbi.nlm.nih.gov/18778909/)]
9. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth Interventions To Support Self-Management In HIV: A Systematic Review. *Open AIDS J* 2017;11:119-132 [FREE Full text] [doi: [10.2174/1874613601711010119](https://doi.org/10.2174/1874613601711010119)] [Medline: [29290888](https://pubmed.ncbi.nlm.nih.gov/29290888/)]
10. Laurence C, Wispelwey E, Flickinger TE, Grabowski M, Waldman AL, Plews-Ogan E, et al. Development of PositiveLinks: A Mobile Phone App to Promote Linkage and Retention in Care for People With HIV. *JMIR Form Res* 2019 Mar 20;3(1):e11578 [FREE Full text] [doi: [10.2196/11578](https://doi.org/10.2196/11578)] [Medline: [30892269](https://pubmed.ncbi.nlm.nih.gov/30892269/)]
11. Dillingham R, Ingersoll K, Flickinger TE, Waldman AL, Grabowski M, Laurence C, et al. PositiveLinks: A Mobile Health Intervention for Retention in HIV Care and Clinical Outcomes with 12-Month Follow-Up. *AIDS Patient Care STDS* 2018 Jun;32(6):241-250 [FREE Full text] [doi: [10.1089/apc.2017.0303](https://doi.org/10.1089/apc.2017.0303)] [Medline: [29851504](https://pubmed.ncbi.nlm.nih.gov/29851504/)]
12. About the Ryan White HIV/AIDS Program. Health Resources and Services Administration. 2019. URL: <https://hab.hrsa.gov/about-ryan-white-hiv-aids-program/about-ryan-white-hiv-aids-program> [accessed 2020-03-03]
13. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci* 2015 Feb 12;10:21 [FREE Full text] [doi: [10.1186/s13012-015-0209-1](https://doi.org/10.1186/s13012-015-0209-1)] [Medline: [25889199](https://pubmed.ncbi.nlm.nih.gov/25889199/)]
14. McNall M, Foster-Fishman PG. Methods of Rapid Evaluation, Assessment, and Appraisal. *American Journal of Evaluation* 2016 Jun 30;28(2):151-168. [doi: [10.1177/1098214007300895](https://doi.org/10.1177/1098214007300895)]
15. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci* 2015 Apr 21;10:53 [FREE Full text] [doi: [10.1186/s13012-015-0242-0](https://doi.org/10.1186/s13012-015-0242-0)] [Medline: [25895742](https://pubmed.ncbi.nlm.nih.gov/25895742/)]
16. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009 Aug 07;4:50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
17. Consolidated Framework for Implementation Science. URL: [cfirguide.org](http://cfirguide.org) [accessed 2020-03-03]
18. Liang S, Kegler MC, Cotter M, Emily P, Beasley D, Hermstad A, et al. Integrating evidence-based practices for increasing cancer screenings in safety net health systems: a multiple case study using the Consolidated Framework for Implementation Research. *Implement Sci* 2016 Aug 02;11:109 [FREE Full text] [doi: [10.1186/s13012-016-0477-4](https://doi.org/10.1186/s13012-016-0477-4)] [Medline: [27485452](https://pubmed.ncbi.nlm.nih.gov/27485452/)]
19. Williams KM, Kirsh S, Aron D, Au D, Helfrich C, Lambert-Kerzner A, et al. Evaluation of the Veterans Health Administration's Specialty Care Transformational Initiatives to Promote Patient-Centered Delivery of Specialty Care: A Mixed-Methods Approach. *Telemed J E Health* 2017 Jul;23(7):577-589. [doi: [10.1089/tmj.2016.0166](https://doi.org/10.1089/tmj.2016.0166)] [Medline: [28177858](https://pubmed.ncbi.nlm.nih.gov/28177858/)]
20. Rosen CS, Matthieu MM, Wiltsey Stirman S, Cook JM, Landes S, Bernardy NC, et al. A Review of Studies on the System-Wide Implementation of Evidence-Based Psychotherapies for Posttraumatic Stress Disorder in the Veterans Health Administration. *Adm Policy Ment Health* 2016 Nov;43(6):957-977. [doi: [10.1007/s10488-016-0755-0](https://doi.org/10.1007/s10488-016-0755-0)] [Medline: [27474040](https://pubmed.ncbi.nlm.nih.gov/27474040/)]
21. Martinez C, Bacigalupe G, Cortada JM, Grandes G, Sanchez A, Pombo H, et al. The implementation of health promotion in primary and community care: a qualitative analysis of the 'Prescribe Vida Saludable' strategy. *BMC Fam Pract* 2017 Feb 17;18(1):23 [FREE Full text] [doi: [10.1186/s12875-017-0584-6](https://doi.org/10.1186/s12875-017-0584-6)] [Medline: [28212606](https://pubmed.ncbi.nlm.nih.gov/28212606/)]
22. Keith RE, Crosson JC, O'Malley AS, Crompton D, Taylor EF. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. *Implement Sci* 2017 Feb 10;12(1):15 [FREE Full text] [doi: [10.1186/s13012-017-0550-7](https://doi.org/10.1186/s13012-017-0550-7)] [Medline: [28187747](https://pubmed.ncbi.nlm.nih.gov/28187747/)]
23. Gimbel S, Rustagi AS, Robinson J, Kouyate S, Coutinho J, Nduati R, with input from the SAIA study team. Evaluation of a Systems Analysis and Improvement Approach to Optimize Prevention of Mother-To-Child Transmission of HIV Using the Consolidated Framework for Implementation Research. *J Acquir Immune Defic Syndr* 2016 Aug 01;72 Suppl 2:S108-S116 [FREE Full text] [doi: [10.1097/QAI.0000000000001055](https://doi.org/10.1097/QAI.0000000000001055)] [Medline: [27355497](https://pubmed.ncbi.nlm.nih.gov/27355497/)]

24. Breimaier HE, Heckemann B, Halfens RJG, Lohrmann C. The Consolidated Framework for Implementation Research (CFIR): a useful theoretical framework for guiding and evaluating a guideline implementation process in a hospital-based nursing practice. *BMC Nurs* 2015;14:43 [FREE Full text] [doi: [10.1186/s12912-015-0088-4](https://doi.org/10.1186/s12912-015-0088-4)] [Medline: [26269693](https://pubmed.ncbi.nlm.nih.gov/26269693/)]
25. Klafke N, Mahler C, von Hagens C, Wensing M, Schneeweiss A, Müller A, et al. How the Consolidated Framework for Implementation Research Can Strengthen Findings and Improve Translation of Research Into Practice: A Case Study. *Oncol Nurs Forum* 2017 Sep 01;44(5):E223-E231. [doi: [10.1188/17.ONF.E223-E231](https://doi.org/10.1188/17.ONF.E223-E231)] [Medline: [28820519](https://pubmed.ncbi.nlm.nih.gov/28820519/)]
26. Damschroder LJ, Lowery JC. Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR). *Implement Sci* 2013 May 10;8:51 [FREE Full text] [doi: [10.1186/1748-5908-8-51](https://doi.org/10.1186/1748-5908-8-51)] [Medline: [23663819](https://pubmed.ncbi.nlm.nih.gov/23663819/)]
27. Luck J, Bowman C, York L, Midboe A, Taylor T, Gale R, et al. Multimethod evaluation of the VA's peer-to-peer Toolkit for patient-centered medical home implementation. *J Gen Intern Med* 2014 Jul;29 Suppl 2:S572-S578 [FREE Full text] [doi: [10.1007/s11606-013-2738-0](https://doi.org/10.1007/s11606-013-2738-0)] [Medline: [24715401](https://pubmed.ncbi.nlm.nih.gov/24715401/)]
28. Haverhals LM, Sayre G, Helfrich CD, Battaglia C, Aron D, Stevenson LD, et al. E-consult implementation: lessons learned using consolidated framework for implementation research. *Am J Manag Care* 2015 Dec 01;21(12):e640-e647 [FREE Full text] [Medline: [26760426](https://pubmed.ncbi.nlm.nih.gov/26760426/)]
29. Garg SK, Lyles CR, Ackerman S, Handley MA, Schillinger D, Gourley G, et al. Qualitative analysis of programmatic initiatives to text patients with mobile devices in resource-limited health systems. *BMC Med Inform Decis Mak* 2016 Feb 06;16:16 [FREE Full text] [doi: [10.1186/s12911-016-0258-7](https://doi.org/10.1186/s12911-016-0258-7)] [Medline: [26851941](https://pubmed.ncbi.nlm.nih.gov/26851941/)]
30. Arrossi S, Paolino M, Orellana L, Thouyaret L, Kohler RE, Viswanath K. Mixed-methods approach to evaluate an mHealth intervention to increase adherence to triage of human papillomavirus-positive women who have performed self-collection (the ATICA study): study protocol for a hybrid type I cluster randomized effectiveness-implementation trial. *Trials* 2019 Feb 26;20(1):148 [FREE Full text] [doi: [10.1186/s13063-019-3229-3](https://doi.org/10.1186/s13063-019-3229-3)] [Medline: [30808379](https://pubmed.ncbi.nlm.nih.gov/30808379/)]
31. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). *Implement Sci* 2016 Oct 26;11(1):146 [FREE Full text] [doi: [10.1186/s13012-016-0510-7](https://doi.org/10.1186/s13012-016-0510-7)] [Medline: [27782832](https://pubmed.ncbi.nlm.nih.gov/27782832/)]
32. Dedoose. URL: <https://www.dedoose.com/> [accessed 2020-03-03]
33. Bardosh KL, Murray M, Khaemba AM, Smillie K, Lester R. Operationalizing mHealth to improve patient care: a qualitative implementation science evaluation of the WeTel texting intervention in Canada and Kenya. *Global Health* 2017 Dec 06;13(1):87 [FREE Full text] [doi: [10.1186/s12992-017-0311-z](https://doi.org/10.1186/s12992-017-0311-z)] [Medline: [29208026](https://pubmed.ncbi.nlm.nih.gov/29208026/)]
34. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. Evaluating the Implementation of a Mobile Phone-Based Telemonitoring Program: Longitudinal Study Guided by the Consolidated Framework for Implementation Research. *JMIR Mhealth Uhealth* 2018 Jul 31;6(7):e10768 [FREE Full text] [doi: [10.2196/10768](https://doi.org/10.2196/10768)] [Medline: [30064970](https://pubmed.ncbi.nlm.nih.gov/30064970/)]
35. Hysong SJ, Smitham KB, Knox M, Johnson K, SoRelle R, Haidet P. Recruiting clinical personnel as research participants: a framework for assessing feasibility. *Implement Sci* 2013 Oct 24;8:125 [FREE Full text] [doi: [10.1186/1748-5908-8-125](https://doi.org/10.1186/1748-5908-8-125)] [Medline: [24153049](https://pubmed.ncbi.nlm.nih.gov/24153049/)]

## Abbreviations

**CFIR:** Consolidated Framework for Implementation Research

**IT:** information technology

**mHealth:** mobile health

**PL:** PositiveLinks

*Edited by L Buis; submitted 06.04.20; peer-reviewed by L Damschroder, L Haverhals, S Farag; comments to author 10.07.20; revised version received 25.09.20; accepted 23.03.21; published 28.04.21.*

*Please cite as:*

Cohn WF, Canan CE, Knight S, Waldman AL, Dillingham R, Ingersoll K, Schexnayder J, Flickinger TE

*An Implementation Strategy to Expand Mobile Health Use in HIV Care Settings: Rapid Evaluation Study Using the Consolidated Framework for Implementation Research*

*JMIR Mhealth Uhealth* 2021;9(4):e19163

URL: <https://mhealth.jmir.org/2021/4/e19163>

doi: [10.2196/19163](https://doi.org/10.2196/19163)

PMID: [33908893](https://pubmed.ncbi.nlm.nih.gov/33908893/)

©Wendy F Cohn, Chelsea E Canan, Sarah Knight, Ava Lena Waldman, Rebecca Dillingham, Karen Ingersoll, Julie Schexnayder, Tabor E Flickinger. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 28.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Using a Mobile Phone App to Analyze the Relationship Between Planned and Performed Physical Activity in University Students: Observational Study

Matthew T Stewart<sup>1</sup>, MSc; Taylor Nezich<sup>1</sup>, MSc; Joyce M Lee<sup>2,3</sup>, MPH, MD; Rebecca E Hasson<sup>1,3</sup>, PhD; Natalie Colabianchi<sup>1,4</sup>, PhD

<sup>1</sup>School of Kinesiology, University of Michigan, Ann Arbor, MI, United States

<sup>2</sup>Susan B. Meister Child Health Evaluation and Research Center, Department of Pediatrics, University of Michigan Medical School, Ann Arbor, MI, United States

<sup>3</sup>Department of Nutritional Sciences, School of Public Health, University of Michigan, Ann Arbor, MI, United States

<sup>4</sup>Institute for Social Research, University of Michigan, Ann Arbor, MI, United States

**Corresponding Author:**

Natalie Colabianchi, PhD

School of Kinesiology

University of Michigan

1402 Washington Heights

Ann Arbor, MI, 48109

United States

Phone: 1 (734) 647 3543

Email: [colabian@umich.edu](mailto:colabian@umich.edu)

## Abstract

**Background:** The relationship between intention and behavior has been well researched, but most studies fail to capture dynamic, time-varying contextual factors. Ecological momentary assessment through mobile phone technology is an innovative method for collecting data in real time, including time-use data. However, only a limited number of studies have examined day-level plans to be physically active and subsequent physical activity behavior using real-time time-use data to better understand this relationship.

**Objective:** This study aims to examine whether plans to be physically active (recorded in advance on an electronic calendar) were associated with objectively assessed physical activity (accelerometry), to identify activities that replaced planned periods of physical activity by using the mobile app Life in a Day (LIAD), and to test the feasibility and acceptability of LIAD for collecting real-time time-use data.

**Methods:** The study included 48 university students who were randomly assigned to 1 of 3 protocols, which were defined by 1, 3, or 5 days of data collection. Participants were asked to record their planned activities on a Google Calendar and were provided with mobile phones with LIAD to complete time-use entries in real time for a set of categories (eg, exercise or sports, eating or cooking, school, or personal care). Participants were instructed to wear an accelerometer on their nondominant wrist during the protocol period. A total of 144 days of protocol data were collected from the 48 participants.

**Results:** Protocol data for 123 days were eligible for analysis. A Fisher exact test showed a statistically significant association between plans and physical activity behavior ( $P=.02$ ). The congruence between plans and behavior was fair (Cohen  $\kappa=0.220$ ; 95% CI 0.028-0.411). Most participants did not plan to be active, which occurred on 75.6% (93/123) of days. Of these 93 days, no physical activity occurred on 76 (81.7%) days, whereas some physical activity occurred on 17 (18.3%) days. On the remaining 24.4% (30/123) of days, some physical activity was planned. Of these 30 days, no physical activity occurred on 18 (60%) days, whereas some physical activity occurred on 12 (40%) days. LIAD data indicated that activities related to screen time most often replaced planned physical activity, whereas unplanned physical activity was often related to active transport. Feasibility analyses indicated little difficulty in using LIAD, and there were no significant differences in feasibility by protocol length.

**Conclusions:** Consistent with previous literature, physical activity plans and physical activity behaviors were linked, but not strongly linked. LIAD offers insight into the relationship between plans and behavior, highlighting the importance of active transport for physical activity and the influence of screen-related behaviors on insufficient physical activity. LIAD is a feasible and practical method for collecting time-use data in real time.

**KEYWORDS**

mobile phone application; physical activity; intention-behavior relationship

## Introduction

### Background

The relationship between physical activity, chronic disease morbidity, and all-cause mortality is well documented. In adults, consistent physical activity can help mitigate risk factors for chronic diseases, including cardiovascular disease, diabetes, cancer, hypertension, obesity, and depression [1]. Although these benefits are widely recognized, approximately 77% of US adults do not meet the current recommendations for aerobic and muscle-strengthening activities [2]. Moreover, marked declines in physical activity participation begin in adolescence, a trend that is sustained in older adulthood [3,4]. Life-transition events, characterized by a significant disruption in previously established routines and habits, can exacerbate these declines [5,6]. One such group experiencing a transitional event is university students, who often cite a newfound lack of time or motivation to be active, subsequently failing to meet the current physical activity guidelines and developing patterns of inactivity that may persist into adulthood [7-9].

Contemporary models suggest that physical activity behavior arises from the interplay of individual, societal, and environmental characteristics [10]. A classic theoretical framework, the theory of planned behavior (TPB), posits that physical activity behavior is predicted by the intention to be active [11,12]. The relationship between intention and behavior has been widely demonstrated in the literature, and as such, the TPB remains a prevalent theoretical framework in physical activity behavior research [13]. However, the strength of the intention-behavior relationship is not consistently robust [14,15]. This inconsistency often stems from the divergence of positive intenders: those who intend to be active but do not act [16]. Research using the TPB has found that those who intend to perform a behavior are more likely to engage in their intended behavior if they make a *plan* to carry out their intentions, including when they will perform their behavior [17].

Recent criticisms of the TPB suggest that the gap between intention and behavior may emerge from the TPB's failure to include the effect of time and context, instead viewing physical activity as a static behavior [18]. Consequently, there is limited evidence regarding *day-level* intentions to be active and subsequent physical activity. A promising method for collecting data in real time is ecological momentary assessment (EMA). This method involves the repeated measurement of current behaviors and contextual factors, often by way of using mobile phones, while minimizing the bias found in methods that require recall well after the activity has occurred [19-21]. The literature regarding EMA has suggested that momentary intentions positively predict subsequent physical activity in the hours following an EMA prompt [22,23]. However, it is not clear whether this is the appropriate time scale, as engaging in physical activity may have logistical issues (eg, finding available time or being able to get to the location where the activity will

be performed), which may need to be considered further in advance. These logistical issues may make having a plan to be active particularly predictive of subsequent physical activity behavior. Furthermore, although the continuous monitoring of EMA has advantages when capturing dynamic, time-varying changes in intention and plans and when predicting behavior, certain EMA features, such as excessive prompting or text messaging, may limit engagement and hinder feasibility [24,25]. When prompting, EMAs often use signal-contingent sampling (sampling based on fixed or random times) or context-aware sampling (sampling based on events defined by automatic sensing technologies), which may interrupt activities when they are occurring and hinder compliance. Conversely, time-use diaries using event-contingent recording may avoid this distraction by asking participants to record information when an event begins and ends [26-28].

### Objectives

Therefore, the purpose of this study is three-fold. The first aim is to examine whether plans to be physically active, recorded in advance on an electronic Google Calendar, were associated with objectively assessed (Actigraph GT9X Link) physical activity over a period of 1, 3, or 5 days. The second aim is to use data from a time-use mobile app (Life in a Day [LIAD]) using event-contingent recording to identify what pursuits occurred when participants had planned to be active but were not and when unplanned physical activity occurred. The final aim is to test the feasibility and acceptability of LIAD for collecting real-time time-use data among university students across 3 different protocol lengths.

## Methods

### Overview

Several recruitment methods were employed. The primary recruitment method was flyers posted in university student spaces detailing information about participation in the study. Researchers also used direct, face-to-face contact with potential participants around campus to aid with recruitment. Participants could inform others of the study, although no compensation was provided to them for their recruitment efforts. Participants recruited via word of mouth from other participants were required to contact us to participate in the study. The subject population was limited to enrolled undergraduates at the University of Michigan aged between 18 years and 25 years. Participants also had to be planning to remain in the area for the 2 weeks following enrollment and be comfortable with using the provided mobile phone and accelerometer. The base compensation was US \$20 for the 1-day protocol, which increased for the 3- and 5-day protocols. Written informed consent was obtained from each participant, detailing their voluntary participation in a study that used mobile devices to collect information regarding time use and physical activity.



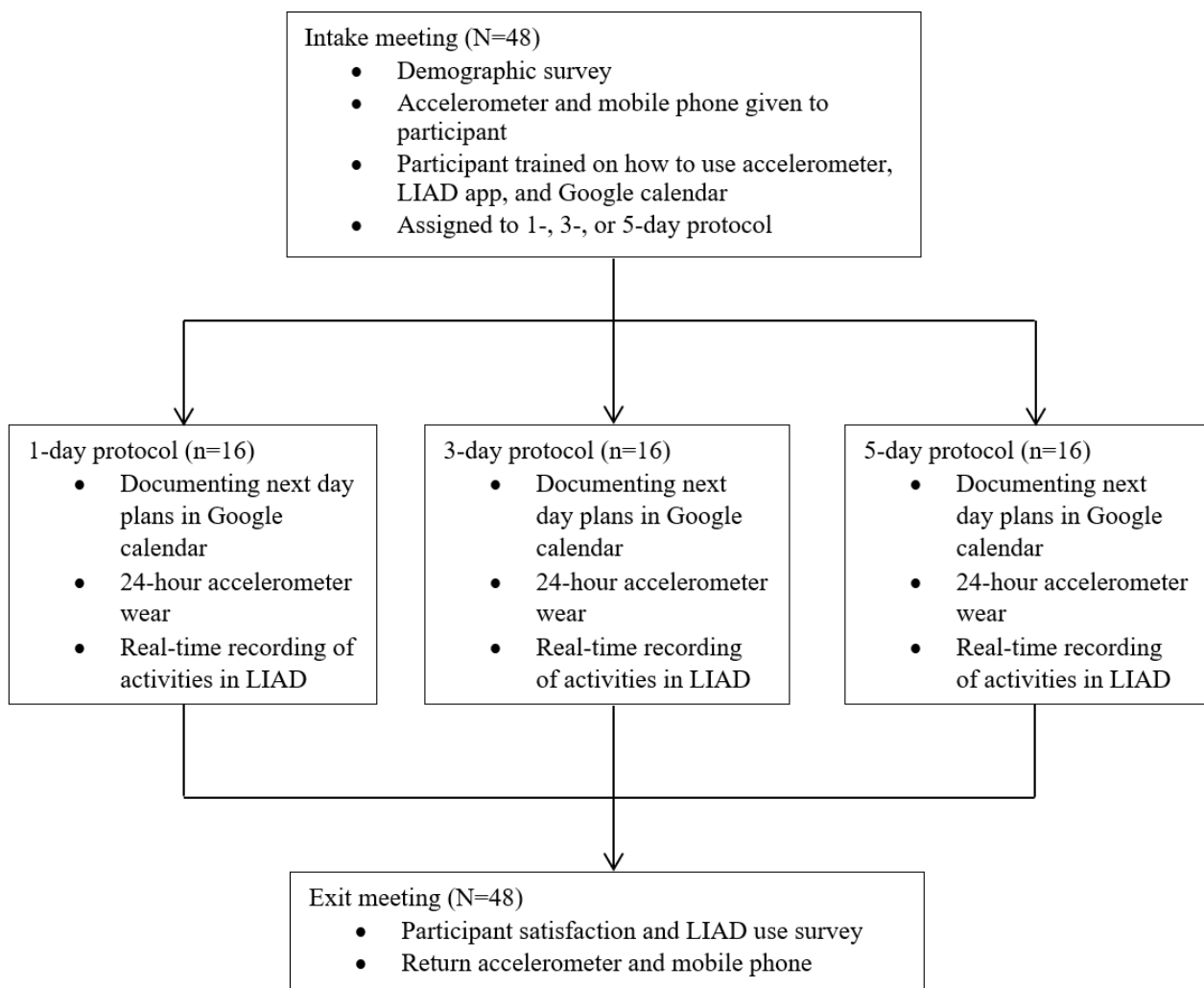
This study was approved by the institutional review board of the University of Michigan.

Data collection began in the summer and ended early in the winter semester, from late July to early February. Breaks in data collection were periodic to avoid scheduled university holidays and examinations, including an extended winter break from early December to mid-January. Participants completed an initial intake meeting, at which they completed an introductory demographic survey consisting of questions regarding age, sex, race, parent education, and engagement in common phone activities (eg, texting, calling, or social media browsing). Participants also received an accelerometer and a mobile phone. The study phone was provided to limit technological issues based on different phone types and potentially limited data availability when using a participant's phone. After providing written informed consent, participants were trained on how to wear the accelerometer and use LIAD (eg, given an explanation of the 12 categories of activities) and were shown how to record physical activity plans and other activities on their Google Calendar. Participants gave verbal confirmation that they understood the protocol and were also given take-home instructions with contact information should they have any

follow-up questions. Participants were randomly assigned to a 1-, 3-, or 5-day protocol at this intake meeting. Three different protocol lengths were used to highlight differences in feasibility. Participants were asked to carry their mobile phone with them and wear the accelerometer for the full duration of the study.

Participants were asked to go about their lives as usual. Before each day of the protocol, participants were asked to record activities planned for the next day on a separate Google Calendar. Throughout each day of the protocol, participants were asked to record activities in real time on LIAD. Activities were filled in retrospectively if they were not completed in real time. At the end of the protocol period, participants returned the study equipment and completed an exit survey, which consisted of questions regarding participant satisfaction, LIAD ease of use and recording in real time, and any issues that may have hindered compliance or engagement. No phone or accelerometer was lost. Once data were downloaded from the mobile phone, the data were transferred to a secure, password-protected server. After transfer, any remaining participant data on the mobile phone were erased. The study flow is shown in [Figure 1](#).

**Figure 1.** Study flow. LIAD: Life in a Day.



## Physical Activity

Physical activity was objectively measured via accelerometry using an Actigraph GT9X Link. Accelerometry is a valid method for capturing physical activity behavior [29,30]. Participants were asked to wear the accelerometer on their nondominant wrist for 24 hours per day for the full protocol period. Each accelerometer was initialized to collect data at 80 Hz (ie, 80 observations per second) for each of the 3 axes. Raw accelerometer data in units of gravity were analyzed using the R package GGIR version 1.8-1, a widely tested software package for quantifying physical activity data across numerous brands of monitors [31-35]. Nonwear was determined by an automated algorithm in GGIR that classifies time as nonwear when the SD is less than 13 mg for 2 of the 3 axes or if the value of each accelerometer axis is less than 150 mg, which is determined using moving windows of 60 minutes within 15-minute increments [32]. A minimum of 10 hours of wear time was required to include that day in the analyses. An acceleration threshold of 200 mg was used to indicate moderate-to-vigorous physical activity (MVPA) based on the work of Hildebrand et al [36]. Data were analyzed using a 10-minute bout metric in which at least 80% of the bout needed to be at or above the MVPA threshold. Minutes of MVPA were calculated for each hour. These files were then merged with Google Calendar and LIAD data in Microsoft Excel to facilitate comparison.

## LIAD App

Activities performed throughout the day were self-tracked using an electronic time-use diary. Time-use diaries are the most common way to assess time use and are a valid method for capturing daily activities [28,37,38]. Electronic methods (eg, web- and app-based modes) were found to produce better quality data and a comparable number of logged activities compared with the paper method [39]. Each participant was provided with a mobile phone (Samsung Galaxy S5) with LIAD downloaded for the duration of the study. LIAD was developed by the Division of Cancer Control and Population Sciences at the US National Cancer Institute in collaboration with MEI Research Ltd and was previously found to be an easy-to-use and acceptable method of measuring time use [40]. The app allowed participants to self-track activities performed throughout the day into 1 of 12 categories: sleeping or doing nothing; personal care; eating or cooking; computer, television, or reading; household activities; entertainment or social activities; school; shopping, appointments, or errands; transport; work (for pay); exercise and sports; and nonsport organized activity. The research team chose these categories based on a review of categories used in the American Time Use Survey, the Panel Study of Income Dynamics (PSID), the Supplement on Disability and Use of Time, and the PSID Childhood Development Supplement [41-43]. If a participant's activity did not fit into one of those categories, participants could add their own category. Participants could also record multiple activities simultaneously (eg, studying while eating dinner). Activities were tracked on LIAD by pressing a *Start* button when the activity began and a *Stop* button when the activity ended (Multimedia Appendices 1-3). Once an activity was recorded, it was automatically added to a calendar built in the app, which participants could view. Though encouraged to

record activities in real time, participants could adjust start and stop times if needed or add activities into gaps within their calendar. Gaps were easily identified by viewing the LIAD calendar, which was accessible via the study mobile phone. LIAD was used to identify the activity that occurred if there was a discrepancy between what was planned on the Google Calendar and what was captured by the accelerometer. For a day to be eligible for analysis, participants needed at least three recorded activities on LIAD, which was considered the minimum number of entries that would occur on a single day.

## Google Calendar

Before each day of the protocol, participants were asked to record their planned activities for the following day on a Google Calendar. Other studies have successfully used calendars to document planned activities, particularly physical activity [44,45]. Unique email addresses and passwords were generated by the research staff and provided to each participant. Instructions were provided to the participants regarding how to fill out the calendar. Participants could either use the web-based Google Calendar or the mobile phone-based app. During data analysis, participants' planned activities were classified into the same categories as those within LIAD to facilitate comparison. If any calendar data were present on a certain day (ie, if the participant recorded at least one planned activity via Google Calendar), that day was included in the analyses as it signified the use of the calendar feature.

## Feasibility

The primary indicators of feasibility included measures of accelerometer wear time, participants' use of LIAD and Google Calendar, and the assessment of any documented participant issues. Accelerometer wear time was measured at the day level for each participant. Participants were asked feasibility questions at the completion of the study, including how easy it was to record activities on LIAD, their estimation of how often they were able to record activities in real time, any issues they may have encountered while recording activities in real time, and willingness to participate in a similar study again.

## Analysis

Demographic characteristics collected from the sample included age, sex (male or female), race (White, Black or African American, Asian, American Indian or Alaskan Native, Native Hawaiian or Pacific Islander, or other), parental education (high school degree, General Educational Diploma, or less; 2-year college or vocational school; 4-year college degree; or graduate degree), and common phone activities (eg, texting, voice calls, video calls, social media browsing, games, maps, directions, and shopping). To assess the congruence between planned and performed physical activity, 4 types of days were calculated: days where physical activity was planned *and* executed, planned *but not* executed, not planned *and* executed, and not planned *or* executed. The Cohen  $\kappa$  and Fisher exact test were used to summarize the congruence between planned and performed physical activity. Participants were categorized as having *planned physical activity* when they had an entry in their Google Calendar that specified an activity that would be physical activity. To determine whether any physical activity was

executed on a given day, we used the accelerometry data, specifically a 10-minute bout of MVPA. Days with no plans for physical activity (according to the Google Calendar) were classified as having no planned physical activity. On days with no plans for physical activity, if any period met our 10-minute threshold for MVPA, we defined it as not planned *and* executed (ie, unplanned physical activity). If a participant planned physical activity on a certain day, we considered that plan to be executed if MVPA occurred at any point during that day for at least 10 minutes, regardless of timing or whether the activity lasted as long as proposed in the calendar. We conducted further analyses to assess the congruent timing of plans and MVPA (ie, executed MVPA that occurred during the same hour as planned in the Google Calendar). With respect to feasibility, compliance with each of the 3 protocol elements was described (accelerometer wear time, Google Calendar entries, and LIAD use). Fisher exact tests were then conducted to determine whether accelerometer wear time, difficulty using LIAD, reporting activities in real time, and completing Google Calendar varied by protocol length. A thematic analysis of each open-ended feasibility question and participant-reported issues was conducted.

## Results

### Overview

In total, there were 48 participants split evenly across the 3 protocol lengths (16 in each). [Table 1](#) presents the demographic characteristics of the sample. For a day to be eligible for analysis, participants needed to have sufficient accelerometry data, LIAD data, and Google Calendar data (sufficient data are defined in the *Methods* section). Of the missing data, 2% (1/48) of participants had no accelerometry data (on the 5-day protocol), expressing during the exit survey that they had forgotten to wear the monitor. Less than 10 hours of valid accelerometer wear time occurred for at least one day for 15% (7/48) of participants. There were missing calendar data on at least one day for 8% (4/48) of participants, all expressing that they had forgotten to complete the calendar the night before. Each participant recorded at least three activities each day on LIAD. Thus, there were missing data across the protocol components for 14.6% (21/144) of days, leaving a total of 85.4% (123/144) of days eligible for analyses.

**Table 1.** Demographic characteristics of the study participants (N=48).

Characteristics	Values
Age (years), mean (SD)	19.8 (1.8)
<b>Sex, n (%)</b>	
Female	37 (77)
Male	11 (23)
<b>Race, n (%)</b>	
White	17 (35)
Black or African American	8 (17)
Asian	19 (40)
Other or multiple races	4 (8)
<b>Parent education: mother, n (%)</b>	
High school, GED <sup>a</sup> , or less	11 (23)
2-year or vocational school	6 (12)
4-year college degree	21 (44)
Graduate degree (eg, Master's, JD <sup>b</sup> , MD <sup>c</sup> , or PhD <sup>d</sup> )	10 (21)
<b>Parent education: father, n (%)</b>	
High school, GED, or less	8 (17)
2-year or vocational school	3 (6)
4-year college degree	17 (35)
Graduate degree (eg, Master's, JD, MD, or PhD)	20 (42)
<b>Common phone activities (multiple responses), n (%)</b>	
Texting	48 (100)
Voice calls	44 (92)
Video calls	27 (56)
Social media browsing	46 (96)
Maps	41 (85)
Directions	42 (88)
Shopping	16 (33)
Games	19 (40)
Other	13 (27)
MVPA <sup>e</sup> minutes, mean (SD)	4.8 (12.3)

<sup>a</sup>GED: General Educational Diploma.

<sup>b</sup>JD: Doctor of Law.

<sup>c</sup>MD: Doctor of Medicine.

<sup>d</sup>PhD: Doctor of Philosophy.

<sup>e</sup>MVPA: moderate-to-vigorous physical activity.

### Planned Versus Performed Physical Activity

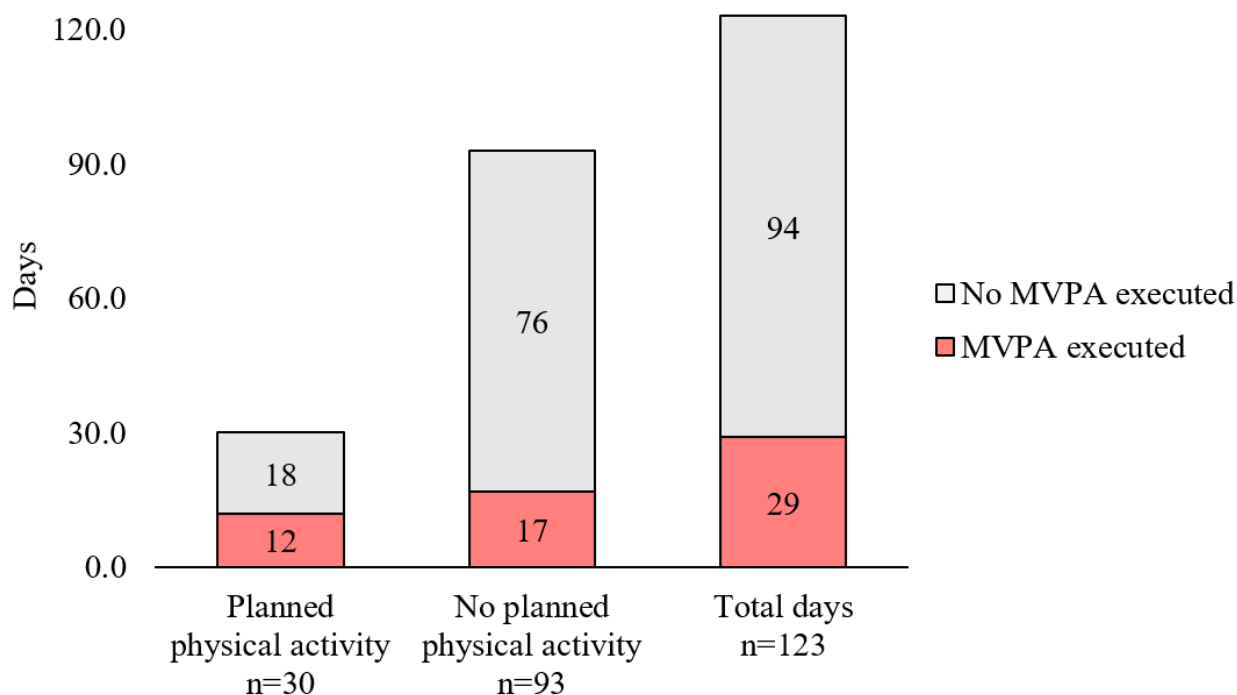
The analysis of planned versus performed physical activity indicated that most participants did not plan any physical activity on their Google Calendars or participate in any physical activity, as shown by accelerometer data. Most participants did not plan to be active, which occurred on 75.6% (93/123) of days. Of those 93 days, no physical activity occurred on 76 (81.7%) days, whereas some physical activity occurred on 17 (18.3%) days.

On the remaining 24.4% (30/123) of days, some physical activity was planned. Of those 30 days, no physical activity occurred on 18 (60%) days, whereas some physical activity occurred on 12 (40%) days. The Fisher exact test showed a statistically significant association between plans and physical activity behavior ( $P=.02$ ). However, when assessing the degree of agreement between plans (or lack of plans) to be active and subsequent behavior using the Cohen  $\kappa$ , only fair agreement was found ( $\kappa=0.220$ ; 95% CI 0.028-0.411). Although the

popular nomenclature for interpreting  $\kappa$  values would suggest fair agreement, it is important to note that the implementation rate of physical activity among those who planned to be active was only 40% [46]. When requiring the accelerometer-derived physical activity to occur during the exact same hour as when

the physical activity was planned, only 25% (3/12) of days had congruent periods of planned physical activity and MVPA, whereas MVPA was performed outside of the planned period on the remaining 75% (9/12) of days. These findings are shown in Figure 2.

**Figure 2.** Planned versus performed physical activity. MVPA: moderate-to-vigorous physical activity.



Data from LIAD were used to reveal activities that replaced planned physical activity or what activities led to unplanned MVPA. The analysis of time-use data from LIAD revealed several endeavors that replaced intended periods of physical activity. These activities were most often related to screen time (eg, using a computer, watching television, or reading), occurring on 56% (10/18) of days. Other categories included social endeavors (eg, spending time with friends or family), occurring on 22% (4/18) of days; sleep, occurring on 11% (2/18) of days; and miscellaneous activities (school and eating), occurring on 11% (2/18) of days. Conversely, for unplanned MVPA, most activities were related to personal care (eg, household chores and getting ready), occurring on 35% (6/17) of days, and active transportation (eg, walking to class or navigating urban areas), occurring on 35% (6/17) of days. Other activities included social endeavors, occurring on 12% (2/17) of days, and miscellaneous activities (physical therapy and classroom-based physical activity), occurring on 18% (3/17) of days. Among those who did not plan to be active and were not active, LIAD data revealed that these participants were mostly involved in work- or school-related activities during the day and screen time activities at night.

### Feasibility

Feasibility analyses indicated an average accelerometer wear time of 17 hours per day. In total, there was insufficient wear time for 8.3% (12/144) of days. All participants in the 1-day protocol had at least 10 hours of wear time per day, whereas

19% (3/16) of participants in the 3-day protocol and 31% (5/16) of participants in the 5-day protocol had invalid wear time for at least one day. This difference was not statistically significant ( $P=.054$ ). When using the Google Calendar, 4% (2/48) of participants in the 1-day protocol and 4% (2/48) of participants in the 5-day protocol were missing calendar data. This difference was not statistically significant ( $P=.53$ ). All participants, regardless of protocol length, recorded at least three activities in the LIAD app. When describing the difficulty of using LIAD, 73% (35/48) of participants mentioned the words *easy* or *simple*. Among those who reported some difficulty, 23% (3/13) of participants were from the 1-day protocol, 54% (7/13) of participants were from the 3-day protocol, and 23% (3/13) of participants were from the 5-day protocol. This difference was not statistically significant ( $P=.22$ ). When using LIAD, 90% (43/48) of participants reported that most of their activities were recorded in real time. Among those who reported only recording some or few of their activities in real time, 40% (2/5) of participants were from the 1-day protocol and 60% (3/5) of participants were from the 3-day protocol. This difference was not statistically significant ( $P=.35$ ). The most commonly reported barriers to record activities in real time included the challenge of remembering to record short activities, trying to start an activity when another is ongoing, and deciding which of the 12 categories to use (they had the option to add their own category). When asked about their participation, 96% (46/48) of participants were willing to enroll in a similar study again. The remaining 4% (2/48) of participants noted a difficulty in

carrying a study phone along with a personal device as a barrier to further participation. Overall, 2869 activities were recorded in the LIAD app across 48 participants (mean 20, SD 9.1 activities per day; range 5-50).

## Discussion

### Principal Findings

Overall, we found fair congruence between plans to be active (or inactive) and physical activity behavior. However, the implementation rate of physical activity among those who planned to be active was only 40%, consistent with previous literature suggesting that there are additional factors that contribute to the relationship between physical activity plans and actual behavior (positive intenders). Intention as a proximal determinant of physical activity behavior is widely theorized, but evidence suggesting intention as a less robust predictor has prompted research into potential augmentations of the TPB model [47,48]. Specifically, dynamic changes in plans to be physically active in real-life settings are not captured well by the traditional TPB [18]. Here, we examined this aspect of the TPB by assessing the relationship between plans for physical activity and subsequent physical activity on a finer time scale (day level). To capture changes in intentions and plans in real time and the contextual factors that may influence this change, methods labeled under the *EMA* umbrella have been developed, assisted by advances in mobile phone technology [49]. The study of mobile phone technology as an innovative arm of *EMA* research has shown promise as a valid, feasible, and powerful tool across multiple populations [50,51].

This study adds unique information to help understand what occurs when plans to be physically active do not materialize. LIAD allows us to answer novel questions, such as those related to identifying replacement activities for planned physical activity and what activities occurred when participants did not plan to be active but were. When examining what activities were displacing physical activity in those periods, we found that most were related to screen time, specifically watching television or working on a computer. The relegation of physical activity in favor of sedentary, screen-related pursuits is consistent with the literature [52,53]. We also found that activities that resulted in MVPA but were not planned as times to be physically active generally included walking to class, navigating the campus, or doing household chores. These findings highlight the potential of using lifestyle choices and active transportation as tools to meet physical activity guidelines, especially when considering that one of the most highly cited reasons for inactivity among university students is the lack of time [8,54-56].

In recognizing these replacement pursuits, we improved our understanding of what factors determine the actual implementation of physical activity at the moment it is planned to occur. Among university students, this motivational flux may occur, in part, due to the accessibility of screen-related pursuits. Greater screen time has been found to be negatively associated with free-time physical activity [57]. Thus, promoting physical activity as part of a lifestyle routine may counteract the motivational flux seen in intention, removing the need to engage in periods of structured physical activity. Future studies could

examine additional contextual factors such as location information, the influence of peers, and types of planned physical activities (eg, running in a park) to help inform the gap between plans to be physically active and actual physical activity and performance of a physical activity that was not planned. Furthermore, LIAD allows for an examination of what activities occurred over the course of the entire day and whether certain activities allow the flexibility to be active as planned, even when it was not during the planned time. Finally, future studies should consider collecting data from individuals over a longer period to understand intraindividual differences in implementation rates and factors that may predict any differences at the individual level.

Although LIAD and similar *EMA* data collection methods can be used as observational tools to shed light on changes in plans across different contexts and settings, they may also serve as a platform to *intervene* in real time. Mobile phone-based *EMA* interventions have previously been found to increase self-awareness of an individual's mood and how they spend their time [49,58,59]. In exit surveys, multiple participants mentioned that LIAD made them more aware of how they spend their time, particularly as it relates to how busy they are. LIAD could be used in an intervention as part of a time management or awareness tool. In turn, this increase in self-awareness and mindfulness can prompt sustained behavioral changes in a range of health behaviors, including physical activity [49,60-62]. Finally, LIAD, in concert with other *EMA* apps, could provide feedback when people follow through on preferred planned behaviors or suggestions and encouragement when their plans fall through.

In this study, LIAD was found to be a feasible method for collecting time-use data in real time over a period of 1-5 days. Our feasibility analyses did not indicate any significant differences highlighted by protocol length, although differences in wear time approached significance. To reliably capture physical activity patterns, 3-5 days of accelerometer wear is recommended [29]. However, studies suggest that longer protocols result in greater sample loss, suggesting a need for researchers to find an appropriate balance [30,63]. With regard to LIAD, we found it to be an overall feasible and practical method of collecting time-use data in real time. Although we encountered few compliance issues, those who failed to record activity in real time were from the 1- or 3-day protocol. Previous literature has suggested that the effect of *EMA* protocol length on compliance, if not small, is somewhat random [64,65]. Rather, the frequency of *EMA* prompting is thought to have a more significant impact on compliance. Some studies have suggested that less frequent prompts ( $\leq 5$  per day) are most beneficial to compliance [64,65]. However, there is limited quantitative evidence regarding the impacts of the frequency of *EMA* prompting on compliance and behavior change [66,67]. In this study, participants were not asked to respond to prompts but rather to record information when events occurred. This may serve as a good time to solicit information, as participants are not in the middle of an activity. This procedure resulted in a high number of recorded events, with up to 50 activities being recorded in a day and an average of 20 activities per day per participant. This, combined with a shorter protocol period (5

days or less), may be responsible for the high compliance with LIAD.

Given the reported difficulty of carrying 2 phones, compliance with reporting activities in real time may be improved if the LIAD app was used on a participant's personal device. Future research should examine the differences between the inclusion of a study phone and the use of a personal device alone, as they relate to compliance and feasibility. We also did not prompt participants in an effort to boost compliance by not interrupting activities that were already occurring. Future research should continue to examine the effect of prompting through mobile devices on feasibility and compliance, including how many prompts are appropriate per day, how often participants should be prompted, and if prompting is more beneficial than allowing participants to record activities at their own volition [64-67]. A balance between recording events as they occur and infrequent prompting may show the greatest benefit to compliance and remembering to record activities in real time [65]. Participants also reported not knowing how to start an activity when another was ongoing as a barrier to record in real time. This may be improved by a more extensive explanation of the LIAD app during the intake meeting and by allowing participants to practice recording activities with the researcher going through various scenarios. Finally, participants noted difficulty in deciding which of the 12 categories to use. Future studies should make use of sequentially nested, fixed lists made possible by electronic time-use surveys, which could facilitate more detailed time-use information and may make the selection process easier for participants [28].

### Strengths and Limitations

This study has several strengths but also a few limitations. It is possible that our bout requirement for physical activity was too restrictive. Most participants recorded no MVPA, including some who planned to be active. Updated physical activity guidelines no longer require activity to be in bouts of 10 minutes or more. However, previous physical activity measurement research has largely used this 10-minute bout metric. Thus, we aimed to maintain consistency with this literature to facilitate comparison with the larger field. We also used wrist-worn accelerometers, although the most common placement for physical activity measurement is at the waist [29,68]. Recent

advancements in triaxial accelerometry and algorithms have improved the accuracy of wrist-worn accelerometers, and using wrist placement has been shown to be less intrusive and increase compliance [69-71]. In this study, we used a well-known method for analyzing raw acceleration data and applying acceleration thresholds for MVPA [31-36]. We also may be limited by the small sample size, which might not be representative of the university student population. Finally, the participants were not required to record their planned activities in Google Calendar using the same categories as the LIAD app. It is possible that the classification of calendar activities into the same categories as the LIAD app by the research team could have introduced measurement errors. A strength of this study is its use of a time-use methodology to examine day-level plans with the possibility of identifying what activities replaced planned physical activity. In naming these physical activity replacements, we provided insight into factors driving the lack of execution of physical activity when it is planned. Other strengths included the ability to compare the feasibility of different lengths of protocol and the use of accelerometer data to objectively capture physical activity.

### Conclusions

In this study of university students, we found fair congruence between plans and physical activity behavior, as captured by a Google Calendar, a real-time time-use mobile phone app, and an accelerometer. A greater understanding of the factors that influence the implementation of physical activity plans may lead to more tailored and effective physical activity interventions that would increase physical activity levels. Most participants reported ease in using LIAD and had limited issues when recording their activities over the course of a day, suggesting that this method of data collection is feasible and practical for use in original research designs or interventions. The increasing popularity of mobile phones as an intervention tool for physical activity behavior indicates that these devices may be a very important entry point when intervening at the individual level [72,73]. With the ever-changing landscape of mobile phone technology and countless hours spent using these devices, future research should continue to examine mobile phones as tools for capturing time-use data, assessing physical activity-related indicators and as a means of intervening to change physical activity behavior.

---

### Acknowledgments

This work was supported by the University of Michigan School of Public Health Momentum Center.

---

### Conflicts of Interest

MTS was supported by grant support under award HL137731. JML was a part of the Medical Advisory Board for GoodRX. During the past 5 years, NC has received grant support from the National Institutes of Health under awards HD101752, HL137731, AG057540, HL132979, NS092706, HL091002, CA188481, CA184478, and CA164137; from the Department of Defense under award 11811976; from the Clark R. Smith Family Foundation; and from the University of Michigan. She also received a speaking fee from the University of Alabama at Birmingham. The remaining authors declare no conflicts of interest.

---

### Multimedia Appendix 1

Participants were instructed to record their activities on the home page of the app, which is located under the "My Activities" page.

[PNG File , 258 KB - [mhealth\\_v9i4e17581\\_app1.png](#) ]

#### Multimedia Appendix 2

Participants can start and stop activities.

[PNG File , 216 KB - [mhealth\\_v9i4e17581\\_app2.png](#) ]

#### Multimedia Appendix 3

A calendar in the upper right corner of the app allows participants to view the activities they have accumulated throughout the days in which they used the Life in a Day app.

[PNG File , 141 KB - [mhealth\\_v9i4e17581\\_app3.png](#) ]

## References

1. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *Can Med Assoc J* 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: [10.1503/cmaj.051351](#)] [Medline: [16534088](#)]
2. Blackwell DL, Clarke TC. State variation in meeting the 2008 federal guidelines for both aerobic and muscle-strengthening activities through leisure-time physical activity among adults aged 18-64: United States, 2010-2015. *Natl Health Stat Report* 2018 Jun(112):1-22 [FREE Full text] [Medline: [30248007](#)]
3. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008 Jan;40(1):181-188. [doi: [10.1249/mss.0b013e31815a51b3](#)] [Medline: [18091006](#)]
4. Sallis JF. Age-related decline in physical activity: a synthesis of human and animal studies. *Med Sci Sports Exerc* 2000;32(9):1598-1600. [doi: [10.1097/00005768-200009000-00012](#)] [Medline: [10994911](#)]
5. Bray SR, Born HA. Transition to university and vigorous physical activity: implications for health and psychological well-being. *J Am Coll Health* 2004;52(4):181-188. [doi: [10.3200/JACH.52.4.181-188](#)] [Medline: [15018429](#)]
6. Allender S, Hutchinson L, Foster C. Life-change events and participation in physical activity: a systematic review. *Health Promot Int* 2008 Jun;23(2):160-172. [doi: [10.1093/heapro/dan012](#)] [Medline: [18364364](#)]
7. Bray SR, Kwan MY. Physical activity is associated with better health and psychological well-being during transition to university life. *J Am Coll Health* 2006;55(2):77-82. [doi: [10.3200/JACH.55.2.77-82](#)] [Medline: [17017303](#)]
8. Arzu D, Tuzun EH, Eker L. Perceived barriers to physical activity in university students. *J Sports Sci Med* 2006;5(4):615-620 [FREE Full text] [Medline: [24357957](#)]
9. Kilpatrick M, Hebert E, Bartholomew J. College students' motivation for physical activity: differentiating men's and women's motives for sport participation and exercise. *J Am Coll Health* 2005;54(2):87-94. [doi: [10.3200/JACH.54.2.87-94](#)] [Medline: [16255320](#)]
10. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. *Annu Rev Public Health* 2006;27:297-322. [doi: [10.1146/annurev.publhealth.27.021405.102100](#)] [Medline: [16533119](#)]
11. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. *Action Control*. Berlin, Heidelberg: Springer; 1985:11-39.
12. Hagger MS, Chatzisarantis NL, Biddle SJ. A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: predictive validity and the contribution of additional variables. *J Sport Exerc Psychol* 2002;24(1):3-32 [FREE Full text] [doi: [10.1123/jsep.24.1.3](#)]
13. Downs DS, Hausenblas HA. The theories of reasoned action and planned behavior applied to exercise: a meta-analytic update. *J Phys Act Health* 2005;2(1):76-97 [FREE Full text] [doi: [10.1123/jpah.2.1.76](#)]
14. Bozionelos G, Bennett P. The theory of planned behaviour as predictor of exercise: the moderating influence of beliefs and personality variables. *J Health Psychol* 1999 Jul;4(4):517-529. [doi: [10.1177/135910539900400406](#)] [Medline: [22021644](#)]
15. de Bruijn GJ. Exercise habit strength, planning and the theory of planned behaviour: an action control approach. *Psychol Sport Exerc* 2011;12(2):106-114 [FREE Full text] [doi: [10.1016/j.psychsport.2010.10.002](#)]
16. Godin G, Conner M. Intention-behavior relationship based on epidemiologic indices: an application to physical activity. *Am J Health Promot* 2008;22(3):180-182. [doi: [10.4278/ajhp.22.3.180](#)] [Medline: [18251118](#)]
17. Gollwitzer PM. Implementation intentions: strong effects of simple plans. *Am Psychol* 1999;54(7):493-503 [FREE Full text] [doi: [10.1037/0003-066x.54.7.493](#)]
18. Dunton GF. Ecological momentary assessment in physical activity research. *Exerc Sport Sci Rev* 2017 Jan;45(1):48-54 [FREE Full text] [doi: [10.1249/JES.0000000000000092](#)] [Medline: [27741022](#)]
19. 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](#)]
20. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. Practical guide to measuring physical activity. *J Acad Nutr Diet* 2014 Feb;114(2):199-208 [FREE Full text] [doi: [10.1016/j.jand.2013.09.018](#)] [Medline: [24290836](#)]
21. 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](#)]



22. Maher JP, Dzibur E, Huh J, Intille S, Dunton GF. Within-day time-varying associations between behavioral cognitions and physical activity in adults. *J Sport Exerc Psychol* 2016 Aug;38(4):423-434. [doi: [10.1123/jsep.2016-0058](https://doi.org/10.1123/jsep.2016-0058)] [Medline: [27634288](https://pubmed.ncbi.nlm.nih.gov/27634288/)]
23. Pickering TA, Huh J, Intille S, Liao Y, Pentz MA, Dunton GF. Physical activity and variation in momentary behavioral cognitions: an ecological momentary assessment study. *J Phys Act Health* 2016 Mar;13(3):344-351. [doi: [10.1123/jpah.2014-0547](https://doi.org/10.1123/jpah.2014-0547)] [Medline: [26284314](https://pubmed.ncbi.nlm.nih.gov/26284314/)]
24. Sternfeld B, Jiang S, Picchi T, Chasan-Taber L, Ainsworth B, Quesenberry CP. Evaluation of a cell phone-based physical activity diary. *Med Sci Sports Exerc* 2012 Mar;44(3):487-495. [doi: [10.1249/MSS.0b013e3182325f45](https://doi.org/10.1249/MSS.0b013e3182325f45)] [Medline: [21857369](https://pubmed.ncbi.nlm.nih.gov/21857369/)]
25. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Smith S. A review of smartphone applications for promoting physical activity. *J Community Med* 2016;2(1):021 [FREE Full text] [Medline: [27034992](https://pubmed.ncbi.nlm.nih.gov/27034992/)]
26. Ruwaard J, Kooistra L, Thong M. Ecological momentary assessment in mental health research: a practical introduction, with examples in r (1st edition - build 2018-11-26). Amsterdam: APH Mental Health; 2018. URL: [https://jruwaard.github.io/aph\\_ema\\_handbook/index.html](https://jruwaard.github.io/aph_ema_handbook/index.html) [accessed 2021-03-18]
27. Aminikhanghahi S, Schmitter-Edgecombe M, Cook DJ. Context-aware delivery of ecological momentary assessment. *IEEE J Biomed Health Inform* 2020 Apr;24(4):1206-1214. [doi: [10.1109/JBHI.2019.2937116](https://doi.org/10.1109/JBHI.2019.2937116)] [Medline: [31443058](https://pubmed.ncbi.nlm.nih.gov/31443058/)]
28. Bauman A, Bittman M, Gershuny J. A short history of time use research; implications for public health. *BMC Public Health* 2019 Jun 03;19(Suppl 2):607 [FREE Full text] [doi: [10.1186/s12889-019-6760-y](https://doi.org/10.1186/s12889-019-6760-y)] [Medline: [31159790](https://pubmed.ncbi.nlm.nih.gov/31159790/)]
29. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005;37(11 Suppl):531-543. [doi: [10.1249/01.mss.0000185657.86065.98](https://doi.org/10.1249/01.mss.0000185657.86065.98)] [Medline: [16294116](https://pubmed.ncbi.nlm.nih.gov/16294116/)]
30. Migueles JH, Cadenas-Sanchez C, Ekelund U, Nyström CD, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med* 2017 Sep;47(9):1821-1845 [FREE Full text] [doi: [10.1007/s40279-017-0716-0](https://doi.org/10.1007/s40279-017-0716-0)] [Medline: [28303543](https://pubmed.ncbi.nlm.nih.gov/28303543/)]
31. Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT. GGIR: a research community-driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *J Measurement Physical Beh* 2019;2(3):188-196. [doi: [10.1123/jmpb.2018-0063](https://doi.org/10.1123/jmpb.2018-0063)]
32. van Hees VT, Gorzelniak L, Dean León EC, Eder M, Pias M, Taherian S, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLoS One* 2013;8(4):- [FREE Full text] [doi: [10.1371/journal.pone.0061691](https://doi.org/10.1371/journal.pone.0061691)] [Medline: [23626718](https://pubmed.ncbi.nlm.nih.gov/23626718/)]
33. Sabia S, van Hees VT, Shipley MJ, Trenell ML, Hagger-Johnson G, Elbaz A, et al. Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors. *Am J Epidemiol* 2014 Mar 15;179(6):781-790 [FREE Full text] [doi: [10.1093/aje/kwt330](https://doi.org/10.1093/aje/kwt330)] [Medline: [24500862](https://pubmed.ncbi.nlm.nih.gov/24500862/)]
34. Rowlands AV, Mirkes EM, Yates TE, Clemes S, Davies M, Khunti K, et al. Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Med Sci Sports Exerc* 2018 Feb;50(2):257-265. [doi: [10.1249/MSS.0000000000001435](https://doi.org/10.1249/MSS.0000000000001435)] [Medline: [28976493](https://pubmed.ncbi.nlm.nih.gov/28976493/)]
35. Rowlands AV, Yates TE, Davies M, Khunti K, Edwardson CL. Raw accelerometer data analysis with GGIR R-package: does accelerometer brand matter? *Med Sci Sports Exerc* 2016 Oct;48(10):1935-1941. [doi: [10.1249/MSS.0000000000000978](https://doi.org/10.1249/MSS.0000000000000978)] [Medline: [27183118](https://pubmed.ncbi.nlm.nih.gov/27183118/)]
36. Hildebrand M, van Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc* 2014 Sep;46(9):1816-1824. [doi: [10.1249/MSS.0000000000000289](https://doi.org/10.1249/MSS.0000000000000289)] [Medline: [24887173](https://pubmed.ncbi.nlm.nih.gov/24887173/)]
37. Kan MY, Pudney S. Measurement error in stylized and diary data on time use. *Socioecol Methodol* 2008;38(1):101-132 [FREE Full text] [doi: [10.1111/j.1467-9531.2008.00197.x](https://doi.org/10.1111/j.1467-9531.2008.00197.x)]
38. Harms T, Gershuny J, Doherty A, Thomas E, Milton K, Foster C. A validation study of the Eurostat harmonised European time use study (HETUS) diary using wearable technology. *BMC Public Health* 2019 Jun 03;19(Suppl 2):455 [FREE Full text] [doi: [10.1186/s12889-019-6761-x](https://doi.org/10.1186/s12889-019-6761-x)] [Medline: [31159770](https://pubmed.ncbi.nlm.nih.gov/31159770/)]
39. Chatzitheochari S, Fisher K, Gilbert E, Calderwood L, Huskinson T, Cleary A, et al. Using new technologies for time diary data collection: instrument design and data quality findings from a mixed-mode pilot survey. *Soc Indic Res* 2018;137(1):379-390 [FREE Full text] [doi: [10.1007/s11205-017-1569-5](https://doi.org/10.1007/s11205-017-1569-5)] [Medline: [29651193](https://pubmed.ncbi.nlm.nih.gov/29651193/)]
40. Ainsworth MC, Pekmezi D, Bowles H, Ehlers D, McAuley E, Courneya KS, et al. Acceptability of a mobile phone app for measuring time use in breast cancer survivors (life in a day): mixed-methods study. *JMIR Cancer* 2018 May 14;4(1):e9 [FREE Full text] [doi: [10.2196/cancer.8951](https://doi.org/10.2196/cancer.8951)] [Medline: [29759953](https://pubmed.ncbi.nlm.nih.gov/29759953/)]
41. American time use survey user's guide. American Time Use Survey.: U.S. Bureau of Labor Statistics; 2019. URL: <https://www.bls.gov/tus/#documents> [accessed 2021-03-18]
42. Freedman VA, Cornman JC. The panel study of income dynamics; supplement on disability and use of time (DUST) user guide: release 2009.1. Institute for Social Research, University of Michigan. 2012. URL: [https://psidonline.isr.umich.edu/DUST/dust09\\_UserGuide.pdf](https://psidonline.isr.umich.edu/DUST/dust09_UserGuide.pdf) [accessed 2021-03-18]
43. Hofferth S, Davis-Kean P, Davis J, Finkelstein J. The child development supplement to the panel study of income dynamics: 1997 user guide. Survey Research Center, Institute for Social Research, University of Michigan. 1997. URL: [https://psidonline.isr.umich.edu/cds/cdsi\\_usergd.pdf](https://psidonline.isr.umich.edu/cds/cdsi_usergd.pdf) [accessed 2021-03-18]

44. Antypas K, Wangberg SC. An internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: short-term results of a randomized controlled trial. *J Med Internet Res* 2014 Mar 11;16(3):e77 [FREE Full text] [doi: [10.2196/jmir.3132](https://doi.org/10.2196/jmir.3132)] [Medline: [24618349](https://pubmed.ncbi.nlm.nih.gov/24618349/)]
45. Robinson SA, Bisson AN, Hughes ML, Ebert J, Lachman ME. Time for change: using implementation intentions to promote physical activity in a randomised pilot trial. *Psychol Health* 2019 Feb;34(2):232-254 [FREE Full text] [doi: [10.1080/08870446.2018.1539487](https://doi.org/10.1080/08870446.2018.1539487)] [Medline: [30596272](https://pubmed.ncbi.nlm.nih.gov/30596272/)]
46. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
47. Rhodes RE, Dickau L. Experimental evidence for the intention-behavior relationship in the physical activity domain: a meta-analysis. *Health Psychol* 2012 Nov;31(6):724-727. [doi: [10.1037/a0027290](https://doi.org/10.1037/a0027290)] [Medline: [22390739](https://pubmed.ncbi.nlm.nih.gov/22390739/)]
48. Rhodes RE, Dickau L. Moderators of the intention-behaviour relationship in the physical activity domain: a systematic review. *Br J Sports Med* 2013;47(4):215-225. [doi: [10.1136/bjsports-2011-090411](https://doi.org/10.1136/bjsports-2011-090411)] [Medline: [22278998](https://pubmed.ncbi.nlm.nih.gov/22278998/)]
49. Runyan JD, Steenbergh TA, Bainbridge C, Daugherty DA, Oke L, Fry BN. A smartphone ecological momentary assessment/intervention "app" for collecting real-time data and promoting self-awareness. *PLoS One* 2013;8(8):e71325 [FREE Full text] [doi: [10.1371/journal.pone.0071325](https://doi.org/10.1371/journal.pone.0071325)] [Medline: [23977016](https://pubmed.ncbi.nlm.nih.gov/23977016/)]
50. Kuntsche E, Labhart F. Using personal cell phones for ecological momentary assessment. *Eur Psychol* 2013;18(1):3-11 [FREE Full text] [doi: [10.1027/1016-9040/a000127](https://doi.org/10.1027/1016-9040/a000127)]
51. Asselbergs J, Ruwaard J, Ejdys M, Schrader N, Sijbrandij M, Riper H. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *J Med Internet Res* 2016 Mar 29;18(3):e72 [FREE Full text] [doi: [10.2196/jmir.5505](https://doi.org/10.2196/jmir.5505)] [Medline: [27025287](https://pubmed.ncbi.nlm.nih.gov/27025287/)]
52. Boone JE, Gordon-Larsen P, Adair LS, Popkin BM. Screen time and physical activity during adolescence: longitudinal effects on obesity in young adulthood. *Int J Behav Nutr Phys Act* 2007 Jun 08;4:26 [FREE Full text] [doi: [10.1186/1479-5868-4-26](https://doi.org/10.1186/1479-5868-4-26)] [Medline: [17559668](https://pubmed.ncbi.nlm.nih.gov/17559668/)]
53. Epstein LH, Roemmich JN, Paluch RA, Raynor HA. Physical activity as a substitute for sedentary behavior in youth. *Ann Behav Med* 2005;29(3):200-209. [doi: [10.1207/s15324796abm2903\\_6](https://doi.org/10.1207/s15324796abm2903_6)] [Medline: [15946114](https://pubmed.ncbi.nlm.nih.gov/15946114/)]
54. Dunn AL, Andersen RE, Jakicic JM. Lifestyle physical activity interventions. History, short- and long-term effects, and recommendations. *Am J Prev Med* 1998 Nov;15(4):398-412. [doi: [10.1016/s0749-3797\(98\)00084-1](https://doi.org/10.1016/s0749-3797(98)00084-1)] [Medline: [9838980](https://pubmed.ncbi.nlm.nih.gov/9838980/)]
55. Dunn AL, Marcus BH, Kampert JB, Garcia ME, Kohl HW, Blair SN. Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness: a randomized trial. *J Am Med Assoc* 1999 Jan 27;281(4):327-334. [doi: [10.1001/jama.281.4.327](https://doi.org/10.1001/jama.281.4.327)] [Medline: [9929085](https://pubmed.ncbi.nlm.nih.gov/9929085/)]
56. Wanner M, Götschi T, Martin-Diener E, Kahlmeier S, Martin BW. Active transport, physical activity, and body weight in adults: a systematic review. *Am J Prev Med* 2012 May;42(5):493-502. [doi: [10.1016/j.amepre.2012.01.030](https://doi.org/10.1016/j.amepre.2012.01.030)] [Medline: [22516490](https://pubmed.ncbi.nlm.nih.gov/22516490/)]
57. Sandercock GH, Ogunleye A, Voss C. Screen time and physical activity in youth: thief of time or lifestyle choice? *J Phys Act Health* 2012;9(7):977-984. [doi: [10.1123/jpah.9.7.977](https://doi.org/10.1123/jpah.9.7.977)] [Medline: [21979868](https://pubmed.ncbi.nlm.nih.gov/21979868/)]
58. Heron KE, Everhart RS, McHale SM, Smyth JM. Using mobile-technology-based Ecological Momentary Assessment (EMA) methods with youth: a systematic review and recommendations. *J Pediatr Psychol* 2017 Nov 01;42(10):1087-1107. [doi: [10.1093/jpepsy/jsx078](https://doi.org/10.1093/jpepsy/jsx078)] [Medline: [28475765](https://pubmed.ncbi.nlm.nih.gov/28475765/)]
59. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 2010 Apr 30;12(2):e10 [FREE Full text] [doi: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)] [Medline: [20439251](https://pubmed.ncbi.nlm.nih.gov/20439251/)]
60. Frates EP, Moore MA, Lopez CN, McMahan GT. Coaching for behavior change in psychiatry. *Am J Phys Med Rehabil* 2011;90(12):1074-1082. [doi: [10.1097/PHM.0b013e31822dea9a](https://doi.org/10.1097/PHM.0b013e31822dea9a)] [Medline: [22019966](https://pubmed.ncbi.nlm.nih.gov/22019966/)]
61. Gilbert D, Waltz J. Mindfulness and health behaviors. *Mindfulness* 2010 Oct 23;1(4):227-234 [FREE Full text] [doi: [10.1007/s12671-010-0032-3](https://doi.org/10.1007/s12671-010-0032-3)]
62. Hudlicka E. Virtual training and coaching of health behavior: example from mindfulness meditation training. *Patient Educ Couns* 2013 Aug;92(2):160-166 [FREE Full text] [doi: [10.1016/j.pec.2013.05.007](https://doi.org/10.1016/j.pec.2013.05.007)] [Medline: [23809167](https://pubmed.ncbi.nlm.nih.gov/23809167/)]
63. Aadland E, Ylvisåker E. Reliability of objectively measured sedentary time and physical activity in adults. *PLoS One* 2015;10(7):e0133296 [FREE Full text] [doi: [10.1371/journal.pone.0133296](https://doi.org/10.1371/journal.pone.0133296)] [Medline: [26192184](https://pubmed.ncbi.nlm.nih.gov/26192184/)]
64. Wen CK, Schneider S, Stone AA, Spruijt-Metz D. Compliance with mobile ecological momentary assessment protocols in children and adolescents: a systematic review and meta-analysis. *J Med Internet Res* 2017 Apr 26;19(4):e132 [FREE Full text] [doi: [10.2196/jmir.6641](https://doi.org/10.2196/jmir.6641)] [Medline: [28446418](https://pubmed.ncbi.nlm.nih.gov/28446418/)]
65. Burke LE, Shiffman S, Music E, Styn MA, Kriska A, Smailagic A, et al. Ecological momentary assessment in behavioral research: addressing technological and human participant challenges. *J Med Internet Res* 2017 Mar 15;19(3):e77 [FREE Full text] [doi: [10.2196/jmir.7138](https://doi.org/10.2196/jmir.7138)] [Medline: [28298264](https://pubmed.ncbi.nlm.nih.gov/28298264/)]
66. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). *J Med Internet Res* 2016 Jun 21;18(6):e151 [FREE Full text] [doi: [10.2196/jmir.4954](https://doi.org/10.2196/jmir.4954)] [Medline: [27328833](https://pubmed.ncbi.nlm.nih.gov/27328833/)]

67. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. *J Med Internet Res* 2009 May 14;11(2):e16 [FREE Full text] [doi: [10.2196/jmir.1138](https://doi.org/10.2196/jmir.1138)] [Medline: [19632970](https://pubmed.ncbi.nlm.nih.gov/19632970/)]
68. Swartz AM, Strath SJ, Bassett DR, O'Brien WL, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc* 2000 Sep;32(9 Suppl):450-456. [doi: [10.1097/00005768-200009001-00003](https://doi.org/10.1097/00005768-200009001-00003)] [Medline: [10993414](https://pubmed.ncbi.nlm.nih.gov/10993414/)]
69. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med* 2014;48(13):1019-1023 [FREE Full text] [doi: [10.1136/bjsports-2014-093546](https://doi.org/10.1136/bjsports-2014-093546)] [Medline: [24782483](https://pubmed.ncbi.nlm.nih.gov/24782483/)]
70. 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](https://doi.org/10.1249/MSS.0000000000001124)] [Medline: [27755355](https://pubmed.ncbi.nlm.nih.gov/27755355/)]
71. 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](https://doi.org/10.1016/j.gaitpost.2015.11.005)] [Medline: [27004628](https://pubmed.ncbi.nlm.nih.gov/27004628/)]
72. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. *J Med Internet Res* 2012 Nov 21;14(6):e161 [FREE Full text] [doi: [10.2196/jmir.2171](https://doi.org/10.2196/jmir.2171)] [Medline: [23171838](https://pubmed.ncbi.nlm.nih.gov/23171838/)]
73. Muntaner A, Vidal-Conti J, Palou P. Increasing physical activity through mobile device interventions: a systematic review. *Health Informatics J* 2016 Sep;22(3):451-469 [FREE Full text] [doi: [10.1177/1460458214567004](https://doi.org/10.1177/1460458214567004)] [Medline: [25649783](https://pubmed.ncbi.nlm.nih.gov/25649783/)]

## Abbreviations

- EMA:** ecological momentary assessment  
**LIAD:** Life in a Day  
**MVPA:** moderate-to-vigorous physical activity  
**PSID:** Panel Study of Income Dynamics  
**TPB:** theory of planned behavior

*Edited by L Buis; submitted 23.12.19; peer-reviewed by Y Liao, N Christina, IC Hou; comments to author 23.03.20; revised version received 14.06.20; accepted 08.03.21; published 29.04.21.*

*Please cite as:*

*Stewart MT, Nezich T, Lee JM, Hasson RE, Colabianchi N*

*Using a Mobile Phone App to Analyze the Relationship Between Planned and Performed Physical Activity in University Students: Observational Study*

*JMIR Mhealth Uhealth* 2021;9(4):e17581

URL: <https://mhealth.jmir.org/2021/4/e17581>

doi: [10.2196/17581](https://doi.org/10.2196/17581)

PMID: [33913812](https://pubmed.ncbi.nlm.nih.gov/33913812/)

©Matthew T Stewart, Taylor Nezich, Joyce M Lee, Rebecca E Hasson, Natalie Colabianchi. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 29.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Self-Efficacy and Motivation to Quit of Smokers Seeking to Quit: Quantitative Assessment of Smoking Cessation Mobile Apps

Nikita B Rajani<sup>1</sup>, MPH; Nikolaos Mastellos<sup>1</sup>, PhD; Filippos T Filippidis<sup>1</sup>, PhD

Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

**Corresponding Author:**

Nikita B Rajani, MPH

Department of Primary Care and Public Health

Imperial College London

School of Public Health, St Dunstan's Road

London, W6 8RP

United Kingdom

Phone: 44 7427615928

Email: [nikita.rajani14@imperial.ac.uk](mailto:nikita.rajani14@imperial.ac.uk)

## Abstract

**Background:** Decreasing trends in the number of individuals accessing face-to-face support are leaving a significant gap in the treatment options for smokers seeking to quit. Face-to-face behavioral support and other interventions attempt to target psychological factors such as the self-efficacy and motivation to quit of smokers, as these factors are associated with an increased likelihood of making quit attempts and successfully quitting. Although digital interventions, such as smoking cessation mobile apps, could provide a promising avenue to bridge the growing treatment gap, little is known about their impact on psychological factors that are vital for smoking cessation.

**Objective:** This study aims to better understand the possible impact of smoking cessation mobile apps on important factors for successful cessation, such as self-efficacy and motivation to quit. Our aim is to assess the self-efficacy and motivation to quit levels of smokers before and after the use of smoking cessation mobile apps.

**Methods:** Smokers seeking to quit were recruited to participate in a 4-week app-based study. After screening, eligible participants were asked to use a mobile app (Kwit or Quit Genius). The smoking self-efficacy questionnaire and the motivation to stop smoking scale were used to measure the self-efficacy and motivation to quit, respectively. Both were assessed at baseline (before app use), midstudy (2 weeks after app use), and end-study (4 weeks after app use). Paired sample two-tailed *t* tests were used to investigate whether differences in self-efficacy and motivation between study time points were statistically significant. Linear regression models investigated associations between change in self-efficacy and change in motivation to quit before and after app use with age, gender, and nicotine dependence.

**Results:** A total of 116 participants completed the study, with the majority being male (71/116, 61.2%), employed (76/116, 65.6%), single (77/116, 66.4%), and highly educated (87/116, 75.0%). A large proportion of participants had a low to moderate dependence on nicotine (107/116, 92.2%). A statistically significant increase of 5.09 points (95% CI 1.83-8.34) from 37.38 points at baseline in self-efficacy was found at the end of the study. Statistically significant increases were also found for the subcomponents of self-efficacy (intrinsic and extrinsic self-efficacies). Similarly, a statistically significant increase of 0.38 points (95% CI 0.06-0.70) from 5.94 points at baseline in motivation to quit was found at the end of the study. Gender, age, and nicotine dependence were not statistically significantly associated with changes in self-efficacy and motivation to quit.

**Conclusions:** The assessed mobile apps positively impacted the self-efficacy and motivation to quit of smokers making quit attempts. This has important implications on the possible future use of digitalized interventions and how they could influence important psychological factors for quitting such as self-efficacy and motivation. However, further research is needed to assess whether digital interventions can supplement or replace traditional forms of therapy.

(*JMIR Mhealth Uhealth* 2021;9(4):e25030) doi:[10.2196/25030](https://doi.org/10.2196/25030)

**KEYWORDS**

smoking cessation; mobile applications; self-efficacy; motivation to quit; mHealth; mobile phone

## Introduction

Smoking is a significant risk factor for many health problems, including lung cancer, heart disease, stroke, and asthma [1]. Although a majority of smokers want to quit smoking in the United Kingdom, the number of smokers making quit attempts has fallen over the past 10 years [2]. Of those individuals trying to quit, the majority are doing so without assistance, although this is associated with lower success rates [3,4]. Among a variety of cessation methods, pharmacological treatment that targets the biological element of addiction, combined with behavioral support, has been found to be the most effective [5]. Some studies have shown that individuals who receive combined behavioral and pharmacological support are 3 times more likely to quit compared with individuals quitting unaided [6]. Behavioral support interventions, particularly individual and group counseling, often use techniques targeting self-efficacy and motivation to quit of smokers, which are factors found to increase the likelihood of smokers making quit attempts and successfully quitting [7-11].

In the context of smoking cessation, self-efficacy is defined as a smoker's confidence in their ability to refrain from smoking when faced with internal (intrinsic self-efficacy) and external stimuli (extrinsic self-efficacy) [12]. Research has found that incorporating self-efficacy into behavioral interventions and integrating it as a central component to the quitting process is effective in treating tobacco use and nicotine dependence [13]. Similarly, enhancing motivation to quit has been found to be a vital part "of the overall treatment for tobacco addiction as it increases smokers' enthusiasm, sense of purpose and will to quit" [14]. Motivation to quit considers the importance placed on quitting and the level of determination a smoker has to quit successfully at a given quit attempt [15].

Despite the effectiveness of face-to-face support targeting important psychological factors such as self-efficacy and motivation, not all individuals are willing or able to access face-to-face support. One study showed that the number of individuals accessing smoking cessation services provided by the National Health Service in the United Kingdom has been continuously declining in part because of budget cuts; a similar trend has been observed in other European countries [16,17]. Further research shows that the proportion of smokers who tried to quit and/or used behavioral support declined in 2017 compared with 2008 [2]. The reduction in the number of individuals accessing face-to-face support leaves a significant gap in the treatment options for smokers seeking to quit. Although some of the alternative cessation methods may also target self-efficacy and motivation to quit, these factors are particularly relevant to digital cessation support, which could be an alternative for face-to-face behavioral treatment.

The provision of digital solutions has gathered increased interest in the field of public health, concurrently with the declining use of behavioral support for smoking cessation. With increased ownership and use of smartphones, smoking cessation interventions delivered via smartphones could be a promising and cost-effective avenue to bridge this gap. One study even reported that in 2015, 400 smoking cessation mobile apps were

available in the various app markets, and this number has most likely risen over the past years [18]. In general, research has found that mobile-based smoking cessation interventions, particularly interventions based on text messages, can positively impact smoking cessation outcomes [19]. Although the evidence for app-based interventions is not as robust, many studies have found positive effects of mobile apps on smoking cessation. For example, Ubhi et al [20] found that smokers who used a cessation app reported higher quitting rates than nonapp users. Despite the proliferation of available cessation apps and their potential to tackle the threat posed by the tobacco epidemic, there is still much to learn about their efficacy and impact on smokers seeking to quit.

Specifically, the literature on whether and how smoking cessation mobile apps can affect psychological success factors such as self-efficacy and motivation to quit is sparse. The few studies that had attempted to investigate this had relatively small sample sizes and/or were purely qualitative [21-23]. Consequently, to address the gap in the literature, this study aims to quantitatively investigate whether the use of smoking cessation mobile apps can positively impact the self-efficacy and motivation to quit of smokers seeking to quit, 2 vital factors for successful cessation.

## Methods

### Study Overview

Smokers seeking to quit were recruited to participate in a 4-week web-based study with no face-to-face contact. Study recruitment and data collection were conducted from June 2019 to July 2020. After an initial screening, eligible participants were asked to use 1 of 2 smoking cessation mobile apps, Kwit or Quit Genius. Participants were asked to complete questionnaires at 3 study time points: baseline (before using the app), midstudy (2 weeks after using the app), and end-study (4 weeks after using the app). A follow-up questionnaire was sent 8 weeks after using the app. Participants were incentivized to participate in the study by providing them with free access to smoking cessation apps with all premium features and the chance to win a £50 (US \$68) Amazon voucher.

### Recruitment

Participants were recruited via social media and posters across public places in London. Interested participants were sent a screening questionnaire to assess their eligibility. Participants who were 18 years or older, proficient in English, current smokers (at least 100 cigarettes smoked in their lifetime and smoked at least one cigarette a day), trying and willing to quit, not using other forms of smoking cessation treatments (including mobile apps), not using or had never used the apps Quit Genius or Kwit, and were not diagnosed with a mental health condition were eligible to participate.

### Sample Size

The research presented in this paper is part of a broader study exploring the use of gamification in smoking cessation mobile apps. Therefore, the sample size was calculated based on a previous study that investigated the impact of gamification elements in a fitness mobile app [24]. With a power level of

$1-\beta=.80$  and a significance level of  $\alpha=.05$ , approximately 112 participants were needed to detect the impact of gamification in the context of the broader study. After accounting for a 20% dropout rate, we aimed to recruit 140 participants and required at least 112 participants to complete the study. Participants completed the study if they self-reported to engage with the app at least once a week for all study weeks and completed all required questionnaires.

### Mobile Apps and App Assignment

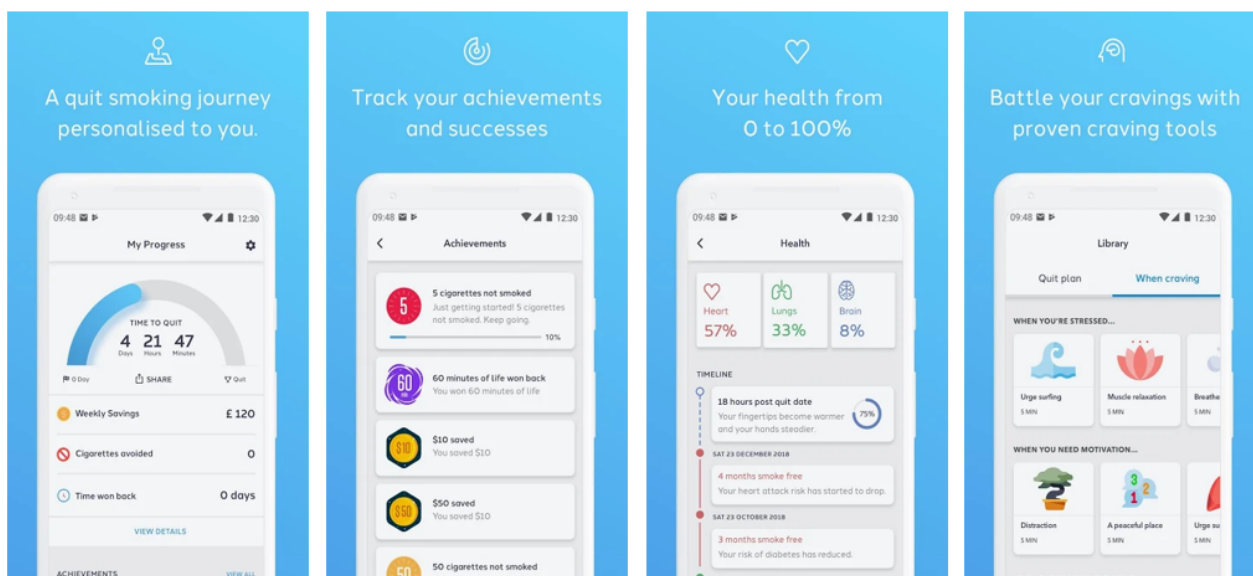
Participants were asked to use 1 of the 2 smoking cessation mobile apps based on their assigned participant identification number (PID). PIDs were assigned on receipt of informed consent and completion of eligibility screening. Participants with even-numbered PIDs were assigned to the mobile app Quit Genius, and participants with odd-numbered PIDs were assigned to the mobile app Kwit. As the analysis presented in this paper is from a broader study investigating the impact of gamification, both apps were chosen based on their embedment of gamification features and adherence to cessation guidelines in the United Kingdom [25]. More specifically, the mobile apps were chosen based on a comprehensive review of smoking cessation mobile apps available on the UK Android and iOS market [25]. The review screened for popular smoking cessation apps based on a minimum rating (ie, rated at least four out of

five) and a minimum number of ratings (ie, 5 individual ratings). After identifying popular apps on the market, the apps were evaluated for their functionalities, adherence to smoking cessation guidelines, and integration of gamification features. The review found that the majority of popular smoking cessation apps had low adherence to evidence-based guidelines. Kwit and Quit Genius were selected because of their high level of adherence to smoking cessation guidelines and adoption of gamification features (relevant for the broader study focusing on gamification).

### Quit Genius

Quit Genius is a gamified smartphone mobile app targeted for smokers seeking to quit smoking and/or maintain their quit status [26]. It delivers personalized behavioral support to individuals based on the principles of cognitive behavioral therapy. The smoking cessation program includes videos, text, and audio recordings to help participants set goals and self-monitor. The app contains a calculator, tracker, cravings toolbox, cigarette diary to log cravings and triggers, and quit coach that provides personalized support. Participants from the study downloaded Quit Genius versions released from June 2019 (v.1.1) to July 2020 (v.1.9). Screenshots of the Quit Genius app are shown in Figure 1.

Figure 1. Screenshots of Quit Genius.

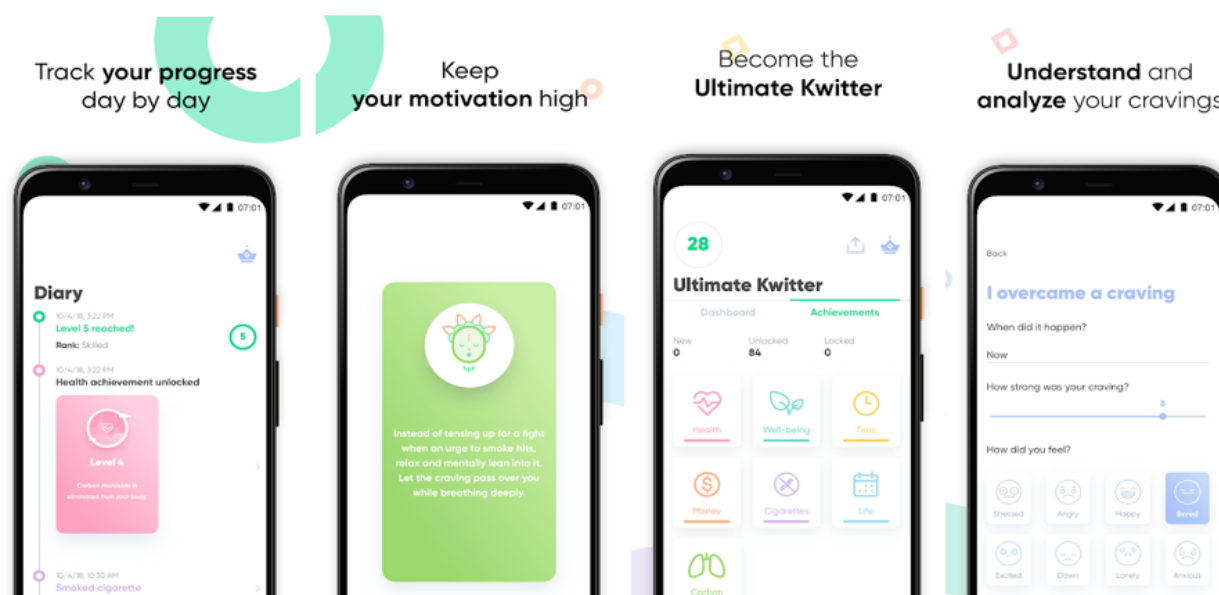


### Kwit

Kwit is a gamified and evidence-based smoking cessation mobile app that helps smokers quit and maintain their quit status [27]. The app is based on cognitive and behavioral therapy principles, gamification, and positive reinforcement. It includes a calculator tracker, motivation cards, social media sharing, and

a smoking diary to log cravings and triggers. Kwit also contains features to help participants deal with relapses and self-monitor their journey to achieve their quitting goals. Participants from the study downloaded Kwit versions released from June 2019 (v.4.1) to July 2020 (v.4.4). Screenshots of the Kwit app are shown in Figure 2.

Figure 2. Screenshots of Kwit.



### Study Time Points

At baseline, participants were asked about general sociodemographic characteristics such as age, gender, marital status, and education. Participants were also asked about their current smoking habits, past quit attempts, self-efficacy, and motivation to quit. After app use, an assessment was conducted at 2 weeks (midstudy) and at 4 weeks (end-study). During both the midstudy and end-study assessments, participants were asked about their self-efficacy and motivation to quit. A follow-up questionnaire was sent at 8 weeks to assess app use, self-efficacy, and motivation to quit.

### Measures

#### *Sociodemographic Factors*

Several sociodemographic factors were assessed at baseline. These included age in years (18-29, 30-41, 42-53, or 54-65), gender (male or female), marital status (single, married, or civil partnered), residence categorized based on World Health Organization regions (Western Pacific, Americas, Southeast Asia, Europe, Africa, and Eastern Mediterranean) [28], education categorized based on the United Nations Educational, Scientific and Cultural Organization's classification (low: primary school completed, medium: secondary school completed, and high: college or university degree) [29], and employment status (unemployed: individuals who are willing and able to work but have no employment; employed; nonemployed: individuals who are unable to work, including students and homemakers).

#### *Nicotine Dependence*

The 6-item Fagerström test was used to measure a participant's tolerance and degree of dependence on nicotine [30]. Participants were classified into levels based on their nicotine dependency score: low (0-4 points), moderate (5-7 points), and high (8-10 points) [30,31].

#### *Past Quit Attempts*

Participants were asked how many serious attempts to stop smoking they made over the past 12 months. Participants were also asked if they used nicotine replacement products, prescribed medications, or mobile apps to help them quit during their past quit attempts in the previous year.

#### *Self-Efficacy*

The smoking self-efficacy questionnaire is a 12-item scale used to measure an individual's confidence in their ability to refrain from smoking on a 5-point Likert scale with the following responses: not at all sure (1), not very sure (2), more or less sure (3), fairly sure (4), and absolutely sure (5) [32]. The total score ranged from 12 to 60, with higher scores indicating greater overall self-efficacy. The 2 subscales include intrinsic and extrinsic self-efficacy, with scores ranging from 6 to 30. Intrinsic self-efficacy refers to the confidence in one's ability to refrain from smoking when faced with internal stimuli such as feeling depressed or anxiety, and extrinsic self-efficacy refers to the confidence in one's ability to refrain from smoking when faced with external stimuli such as being with other smokers and drinking alcohol.

#### *Motivation to Quit*

Motivation to quit smoking is the level of importance a smoker places on quitting and the level of determination a smoker has to quit successfully at a given quit attempt [15]. Motivation to quit smoking was measured using a 2-item measure frequently used in past smoking cessation studies [33-35]. Participants were asked, "How important is it to you to give up smoking altogether at this attempt?" Responses included "desperately important," "very important," "quite important," and "not all that important." Participants were also asked, "How determined are you to give up smoking at this attempt?" Responses included "extremely determined," "very determined," "quite determined," and "not all that determined." Responses for both questions were coded and totaled, resulting in a range from 2 to 8, with a higher score indicating a higher level of overall motivation.

### Statistical Analysis

The analyses were conducted using STATA 13.1 (StataCorp). Descriptive statistics were used to present general participant characteristics, including current smoking habits and information on past quit attempts. Mean values of self-efficacy and motivation to quit were calculated at baseline, midstudy, and end-study. Owing to a low response rate of 40% to the follow-up questionnaire at 8 weeks, data from this questionnaire are not presented. Two-tailed paired sample *t* tests were used to test whether differences in self-efficacy and motivation to quit at various time points of the study were statistically significant. In addition, we explored various linear regression models to examine factors associated with changes in self-efficacy and motivation to quit. On the basis of an iterative process considering the fit of the model with the data, our linear regression models investigated the association between change in self-efficacy and change in motivation to quit before and after

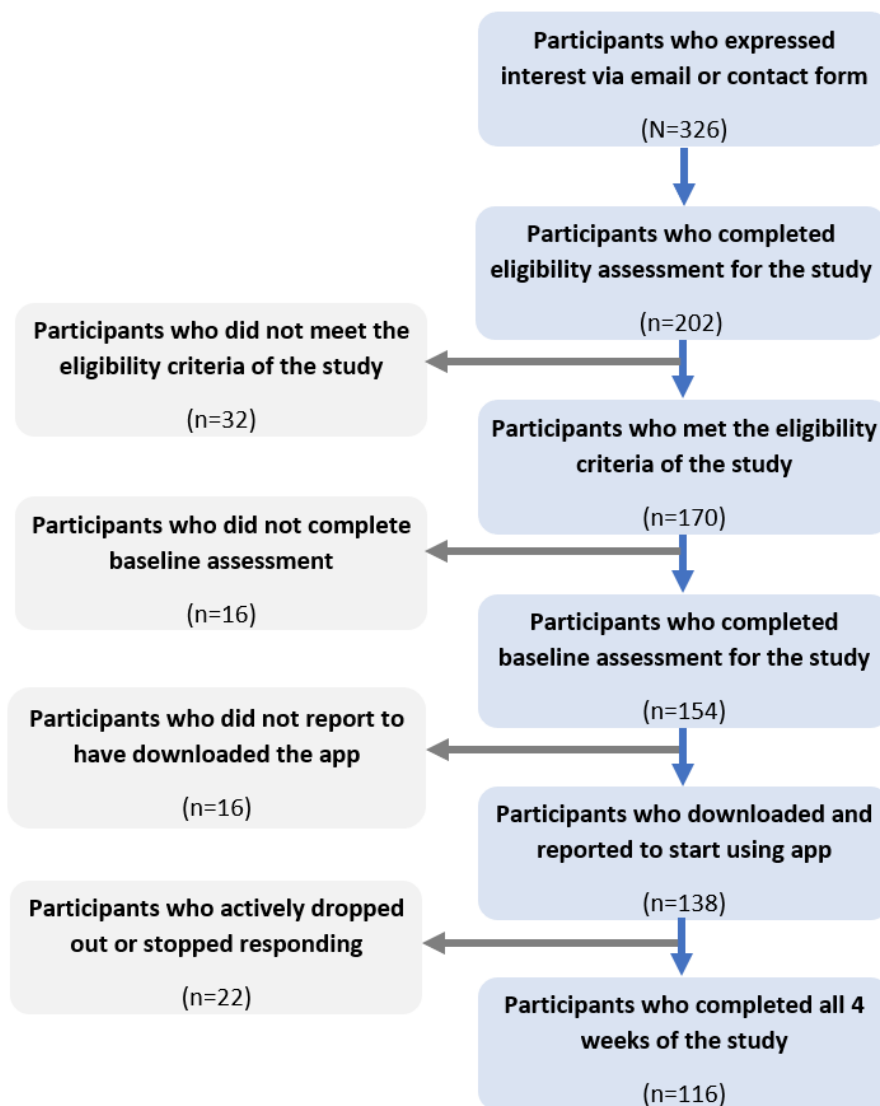
app use with age, gender, and nicotine dependence. Statistical significance was determined at 5% ( $P=.05$ ), and 95% CIs were reported for all coefficients presented.

## Results

### Study Participants

The flowchart in Figure 3 displays the number of individuals completing the different stages of the study. Among the 326 individuals who expressed interest, 202 (62.0%) completed the eligibility questionnaire. Of the 202, 170 (84.2%) met the eligibility criteria. Of the 154 participants who completed the baseline assessment and were sent app installation instructions, 138 (89.6%) self-reported that they successfully installed and logged on to the assigned app. In total, 116 participants completed the entire study, and their characteristics are shown in Table 1.

**Figure 3.** Overview of participant numbers from expression of interest to study completion.





**Table 1.** General characteristics, smoking habits, and past use of quitting aids (n=116).

General characteristics	Respondents, n (%)
<b>Age (years)</b>	
18-29	49 (42.2)
30-41	41 (35.3)
42-53	15 (12.9)
54-65	11 (9.5)
<b>Gender</b>	
Male	71 (61.2)
Female	45 (38.8)
<b>Education</b>	
Low (primary school)	8 (6.9)
Medium (secondary school)	21 (18.1)
High (university or college degree)	87 (75.0)
<b>Marital status</b>	
Single	77 (66.4)
Married or civil partnered	39 (33.6)
<b>Employment status</b>	
Employed	76 (65.5)
Nonemployed	31 (26.7)
Unemployed	6 (5.2)
Prefer not to answer	3 (2.6)
<b>World Health Organization regions</b>	
Western Pacific	4 (3.4)
Americas	10 (8.6)
Southeast Asia	16 (13.8)
Europe	67 (57.8)
Africa	17 (14.7)
Eastern Mediterranean	2 (1.7)
<b>Daily smoking (number of cigarettes)</b>	
≤10	63 (54.3)
11-20	43 (37.1)
21-30	8 (6.9)
≥31	2 (1.7)
<b>Fagerström nicotine dependence</b>	
Low (0-4)	62 (53.4)
Moderate (5-7)	45 (38.8)
High (8-10)	9 (7.8)
<b>Age (years) started smoking</b>	
<15	30 (25.9)
16-29	82 (70.7)
≥30	4 (3.4)
<b>Past use of nicotine replacement therapy to quit</b>	
No	64 (55.2)

General characteristics	Respondents, n (%)
Yes	52 (44.8)
<b>Past use of cessation medication</b>	
No	95 (81.9)
Yes	21 (18.1)
<b>Past use of mobile apps to quit</b>	
No	108 (93.1)
Yes	8 (6.9)

Table 1 shows that the majority of participants were male (71/116, 61.2%), single (77/116, 66.4%), employed (76/116, 65.5%), and reported having a high level of education (87/116, 75.0%). With regard to smoking habits, more than half of the participants reported that they smoked 10 or fewer cigarettes daily (63/116, 54.3%). Similarly, the majority of respondents had low or moderate dependence on nicotine according to the Fagerström test (107/116, 92.2%) and reported that they did not use nicotine replacement treatments (64/116, 55.2%), medications (95/116, 81.9%), or mobile apps (108/116, 93.1%) in the past 12 months to aid with quit attempts.

### Self-Efficacy and Motivation to Quit

Table 2 displays the participants' self-efficacy at baseline, midstudy, and end-study. After 4 weeks of using the app (end-study), intrinsic, extrinsic, and overall self-efficacy increased by 2.7, 2.4, and 5.1 points, respectively, compared with baseline values of 18.4, 19.0, and 37.4, respectively. Two-tailed paired sample *t* tests showed that these increases were statistically significant. Similarly, perceived importance, determination, and overall motivation to quit increased by 0.2, 0.2, and 0.4 points, respectively, compared with baseline values of 3, 2.9, and 5.9. These increases were statistically significant. However, the mean differences in self-efficacy and motivation to quit between midstudy and end-study were not statistically significant.

**Table 2.** Mean self-efficacy and motivation to quit scores across different study time points (n=116).

Characteristics	Baseline, mean (SD)	Midstudy, mean (SD)	End-study, mean (SD)	Baseline versus mid-study, mean difference (95% CI)	Baseline versus end-study, mean difference (95% CI)	Midstudy versus end-study, mean difference (95% CI)
<b>Self-efficacy</b>						
Intrinsic (6-30)	18.4 (7.2)	20.6 (5.6)	21.0 (6.1)	2.2 (0.8 to 3.7)	2.7 (1.0 to 4.4)	0.5 (-0.7 to 1.6)
Extrinsic (6-30)	19.0 (7.0)	20.8 (6.2)	21.4 (6.3)	1.8 (0.2 to 3.4)	2.4 (0.7 to 4.1)	0.6 (-0.6 to 1.8)
Overall (12-60)	37.4 (13.3)	41.4 (10.5)	42.5 (11.5)	4.0 (1.2 to 6.8)	5.1 (1.8 to 8.3)	1.1 (-1.0 to 3.2)
<b>Motivation</b>						
Importance (1-4)	3.0 (0.8)	3.2 (0.8)	3.2 (0.8)	0.2 (0.0 to 0.3)	0.2 (0.0 to 0.4)	0.1 (-0.1 to 0.2)
Determination (1-4)	2.9 (0.8)	3.0 (0.9)	3.1 (0.9)	0.1 (-0.1 to 0.3)	0.2 (0.0 to 0.4)	0.1 (-0.1 to 0.2)
Overall (2-8)	5.9 (1.4)	6.2 (1.5)	6.3 (1.7)	0.2 (0.0 to 0.5)	0.4 (0.1 to 0.7)	0.1 (-0.1 to 0.3)

The results of the linear regression models are shown in Table 3. Our analysis found that age was not statistically associated with a change in self-efficacy between end-study and baseline ( $\beta=-.07$ , 95% CI  $-.39$  to  $.24$ ). We also found that female participants self-reported a smaller change in self-efficacy compared with males; however, this result was also not statistically significant ( $\beta=-.91$ , 95% CI  $-8.09$  to  $6.26$ ).

Moreover, we found no association between a participant's level of nicotine dependence and the change in overall self-efficacy. Similar to self-efficacy, our analysis found no statistically significant associations between age, gender, and nicotine dependence with change in overall motivation to quit between baseline and end-study.

**Table 3.** Linear regression models examining factors associated with change in self-efficacy and change in motivation to quit between end-study and baseline (n=116).

Variables	Change in overall self-efficacy (end-study vs baseline), $\beta$ coefficient (95% CI)	Change in overall motivation (end-study vs baseline), $\beta$ coefficient (95% CI)
Age (years)	-.07 (-.39 to .24)	-.02 (-.05 to .01)
<b>Gender</b>		
Male	Reference	Reference
Female	-.91 (-8.08 to 6.26)	.17 (-.54 to .87)
<b>Nicotine dependence</b>		
Low	Reference	Reference
Moderate	-.01 (-7.07 to 7.04)	-.11 (-.81 to .58)
High	1.87 (-10.99 to 14.74)	.13 (-1.14 to 1.40)
Constant	7.82 (-3.33 to 18.97)	.96 (-.14 to 2.06)

## Discussion

### Principal Findings

We found that self-reported self-efficacy and motivation to quit among participants using the assigned apps increased after app use. However, this increase largely occurred during the first 2 weeks of app use and then plateaued. We also found that age, gender, and nicotine dependence were not associated with changes in overall self-efficacy and motivation to quit between end-study and baseline.

The observed statistically significant increase in overall self-efficacy from baseline to midstudy and baseline to end-study implies that smokers seeking to quit smoking experienced an increase in perceived confidence in their ability to refrain from smoking. The same was found when examining intrinsic and extrinsic self-efficacy, suggesting that participants experienced increased confidence in their ability to refrain from smoking when faced not only with internal stimuli such as their feelings and cravings but also with external stimuli such as socializing with other smokers and drinking alcohol. This finding is important for smokers seeking to quit, as several studies have shown that high self-efficacy is associated with better smoking cessation outcomes [7-10].

The increase in self-efficacy after app use was generally consistent with previous studies. For example, one study examining the impact of a mobile app promoting smoking cessation in hospitalized patients found positive changes in self-efficacy among patients after app use [23]. However, pre- and post self-efficacy scores were only 45 minutes apart. Similarly, Hoepfner et al [21] found that among 30 participants, higher self-efficacy was reported 2 weeks after using a smoking cessation mobile app compared with baseline.

We also noted that the majority of the increase in self-efficacy was found between baseline and midstudy, after which self-efficacy levels stabilized. This could mean that the mobile apps may have a saturated effect by the midstudy point; therefore, the first 2 weeks act as a *ramp-up* phase, after which self-efficacy plateaus. Boardman et al [36] conducted a study among 600 African American smokers and found that the smokers continued to sustain high levels of self-efficacy and

motivation scores up to 6 months post intervention. Although our findings are consistent with the findings from the study by Boardman et al [36], this study did not investigate the long-term maintenance of self-efficacy levels after mobile app use; therefore, this could be explored in future research.

Similar to self-efficacy, a statistically significant difference in determination, importance, and overall motivation to quit was found between baseline and end-study. This shows that participants experienced an increase in their perceived determination to quit smoking and how important they felt it was to quit at this attempt after using the app compared with baseline. Both apps contained several features that could have led to increased motivation. For example, possible features embedded in the Kwit app that could have led to higher motivation include unlocking achievements, motivation cards, trackers, and craving management tools. On the other hand, Quit Genius includes features such as achievement badges, cravings management, and a quit coach. Comparing the 2 apps and/or understanding which app-specific features were associated with changes in self-efficacy and motivation to quit was beyond the scope of our analysis.

The statistically significant increase in motivation to quit found in our analysis is in line with another study investigating the impact of Quit Genius, which also reported that participants had higher motivation to quit after using the app [22]. However, that study was purely qualitative, with only 15 participants using the app for 1 week. Similarly, Hoepfner et al [21] found increased motivation to quit levels among smokers prescribed to use a cessation app for 3 weeks. It is interesting to note that, on average, participants who enrolled in the study had relatively high motivation to quit at baseline, which could be because of participants with high motivation self-selecting into the study [37].

Furthermore, we did not find a statistically significant association between age, gender, and nicotine dependence with change in overall self-efficacy and motivation to quit. This suggests that the mobile apps had a similar effect on participants' self-efficacy and motivation to quit at the end of the study compared with baseline, regardless of age, gender, and level of nicotine dependence. However, this might not be

generalizable, as our sample had a majority of male, educated, and employed participants with low to moderate nicotine dependence. It may also have been underpowered to detect nuanced differences in effectiveness. A more diverse sample could allow for stronger inference on benefits of these apps across users of all demographics.

The positive impact of the mobile apps on the self-efficacy and motivation to quit of smokers highlights the importance and role of these psychological factors during a quit attempt. According to Elshatarat et al [13], self-efficacy is “a central concept of the quitting process” and is vital for achieving cessation and preventing relapse. Self-efficacy beliefs can determine how and when a smoker will initiate their coping behaviors and how long they are sustained when experiencing cravings or withdrawal symptoms [13]. Similarly, theories surrounding motivation to quit suggest that it is also a critical ingredient in quitting and plays an important role in the quitting process by enhancing a smoker’s intention to quit [11]. Studies have also found that apps can influence other psychological factors, such as well-being and psychological empowerment, to help smokers quit. For example, Lin et al [38] investigated the impact of Quit Genius on smokers seeking to quit and found that the app was able to enhance hedonic well-being and psychologically empower smokers, which, in turn, significantly increased the odds of successfully quitting.

Future research should continue to build upon our understanding of how mobile app solutions can positively impact psychological factors, such as self-efficacy, motivation to quit, and empowerment, which are found to be vital for successful cessation. For example, certain features or design elements may be more effective than others in influencing important psychological factors, promoting health behavior change, and improving quit rates. App developers and tobacco cessation and behavior change specialists could benefit from working together to develop effective digital cessation programs that contain features targeted at improving and enhancing psychological factors that may play a role in the quitting process. Finally, as past studies have shown that internet-based interventions can help the disadvantaged more, the possibility of providing effective digitalized interventions could help reduce health inequalities [39,40]. Consequently, more research that rigorously assesses the use and impact of digitalized behavioral

interventions could be very valuable for public health policy makers working to attenuate health disparities.

### Limitations

One of the limitations of our study is that the majority of participants had low dependence on nicotine, which could affect baseline self-efficacy and/or motivation to quit. Future research could replicate the research on high-dependency smokers to see whether the results are generalizable. Similarly, unlike our study, future studies could also include participants with mental health conditions to ensure that the findings are generalizable to this population subgroup as well. Another limitation is the reliance on only self-reported data, which can lead to biases such as social desirability bias and, therefore, may not always be the most reliable. Moreover, our study had some methodological limitations. For example, participants were assigned to 1 of 2 mobile apps to ensure accurate data collection before and after app use. However, in reality, smokers would naturally self-select interventions on their app stores. It could also be that our study consisted of individuals with higher motivation than the general population, causing some volunteer bias. Finally, not enough follow-up data were collected because of low response rates at 8 weeks to be able to comment on long-term impact. Therefore, future studies could investigate whether the effects of mobile apps are sustained in the long term and how this can be compared with face-to-face behavioral cessation programs. Despite these limitations, this study develops a better understanding of the impact of smoking cessation apps and could provide a basis for future randomized controlled trials.

### Conclusions

In conclusion, we found that smoking cessation mobile apps could have a positive impact on important psychological factors associated with better cessation outcomes. This has important implications for the development and use of mobile apps as evidence-based support for smoking cessation. Although this research might provide insights for the development of future apps, further research is required to enhance our understanding of how digitalized interventions could positively impact self-efficacy and other psychological factors vital for successful cessation. The limitations of our study methodology highlight the issues that future research can address differently. More rigorous and evidence-based research is vital to determine whether digital interventions can supplement or replace traditional forms of therapy.

---

### Acknowledgments

The study was conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki 1964, and later versions. Ethical approval was obtained from the Joint Research Imperial College London Research Ethics Committee before the beginning of the study (reference 19IC5158). This research did not receive a grant or funding from any agencies that are public, commercial, or within the not-for-profit sector. Quit Genius and Kwit provided free access to the study participants but had no other financial or material input.

---

### Authors' Contributions

NBR, NM, and FTF worked together to design the study and develop the study protocol. NBR acted as the study coordinator to collect the necessary data and handle the day-to-day management of the study. NBR performed the statistical analysis with

guidance provided by NM and FTF. NM and FTF reviewed and revised the manuscript. All authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

## References

1. Hackshaw A, Morris JK, Boniface S, Tang J, Milenković D. Low cigarette consumption and risk of coronary heart disease and stroke: meta-analysis of 141 cohort studies in 55 study reports. *Br Med J* 2018 Jan 24;360:j5855 [FREE Full text] [doi: [10.1136/bmj.j5855](https://doi.org/10.1136/bmj.j5855)] [Medline: [29367388](https://pubmed.ncbi.nlm.nih.gov/29367388/)]
2. Garnett C, Tombor I, Beard E, Jackson S, West R, Brown J. Changes in smoker characteristics in England between 2008 and 2017. *Addiction* 2020 Apr;115(4):748-756 [FREE Full text] [doi: [10.1111/add.14882](https://doi.org/10.1111/add.14882)] [Medline: [31914486](https://pubmed.ncbi.nlm.nih.gov/31914486/)]
3. Bauld L, Hiscock R, Dobbie F, Aveyard P, Coleman T, Leonardi-Bee J, et al. English Stop-Smoking Services: One-Year Outcomes. *Int J Environ Res Public Health* 2016 Nov 24;13(12):1175 [FREE Full text] [doi: [10.3390/ijerph13121175](https://doi.org/10.3390/ijerph13121175)] [Medline: [27886140](https://pubmed.ncbi.nlm.nih.gov/27886140/)]
4. Hughes J, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* 2004 Jan;99(1):29-38. [doi: [10.1111/j.1360-0443.2004.00540.x](https://doi.org/10.1111/j.1360-0443.2004.00540.x)] [Medline: [14678060](https://pubmed.ncbi.nlm.nih.gov/14678060/)]
5. Stead L, Koilpillai P, Fanshawe T, Lancaster T. Combined pharmacotherapy and behavioural interventions for smoking cessation. *Cochrane Database Syst Rev* 2016 Mar 24;3:1-98. [doi: [10.1002/14651858.CD008286.pub3](https://doi.org/10.1002/14651858.CD008286.pub3)] [Medline: [27009521](https://pubmed.ncbi.nlm.nih.gov/27009521/)]
6. Kotz D, Brown J, West R. 'Real-world' effectiveness of smoking cessation treatments: a population study. *Addiction* 2014 Mar 20;109(3):491-499. [doi: [10.1111/add.12429](https://doi.org/10.1111/add.12429)] [Medline: [24372901](https://pubmed.ncbi.nlm.nih.gov/24372901/)]
7. Smit ES, Hoving C, Schelleman-Offermans K, West R, de Vries H. Predictors of successful and unsuccessful quit attempts among smokers motivated to quit. *Addict Behav* 2014 Sep;39(9):1318-1324. [doi: [10.1016/j.addbeh.2014.04.017](https://doi.org/10.1016/j.addbeh.2014.04.017)] [Medline: [24837754](https://pubmed.ncbi.nlm.nih.gov/24837754/)]
8. Ockene J, Emmons K, Mermelstein R, Perkins KA, Bonollo DS, Voorhees CC, et al. Relapse and maintenance issues for smoking cessation. *Health Psychol* 2000 Jan;19(1S):17-31 [FREE Full text] [doi: [10.1037/0278-6133.19.suppl1.17](https://doi.org/10.1037/0278-6133.19.suppl1.17)] [Medline: [10709945](https://pubmed.ncbi.nlm.nih.gov/10709945/)]
9. Li L, Borland R, Yong H, Fong GT, Bansal-Travers M, Quah ACK, et al. Predictors of smoking cessation among adult smokers in Malaysia and Thailand: findings from the International Tobacco Control Southeast Asia Survey. *Nicotine Tob Res* 2010 Oct 30;12(Supplement 1):S34-S44 [FREE Full text] [doi: [10.1093/ntr/ntq030](https://doi.org/10.1093/ntr/ntq030)] [Medline: [20889478](https://pubmed.ncbi.nlm.nih.gov/20889478/)]
10. Osler M, Prescott E. Psychosocial, behavioural, and health determinants of successful smoking cessation: a longitudinal study of Danish adults. *Tob Control* 1998 Sep 01;7(3):262-267 [FREE Full text] [doi: [10.1136/tc.7.3.262](https://doi.org/10.1136/tc.7.3.262)] [Medline: [9825421](https://pubmed.ncbi.nlm.nih.gov/9825421/)]
11. Piñeiro B, López-Durán A, Del Río EF, Martínez Ú, Brandon TH, Becoña E. Motivation to quit as a predictor of smoking cessation and abstinence maintenance among treated Spanish smokers. *Addict Behav* 2016 Feb;53:40-45. [doi: [10.1016/j.addbeh.2015.09.017](https://doi.org/10.1016/j.addbeh.2015.09.017)] [Medline: [26441045](https://pubmed.ncbi.nlm.nih.gov/26441045/)]
12. Etter J, Bergman M, Humair J, Perneger T. Development and validation of a scale measuring self-efficacy of current and former smokers. *Addiction* 2000 Jun;95(6):901-913. [doi: [10.1046/j.1360-0443.2000.9569017.x](https://doi.org/10.1046/j.1360-0443.2000.9569017.x)] [Medline: [10946439](https://pubmed.ncbi.nlm.nih.gov/10946439/)]
13. Elshatarat RA, Yacoub MI, Khraim FM, Saleh ZT, Afaneh TR. Self-efficacy in treating tobacco use: A review article. *Proceedings of Singapore Healthcare* 2016 Sep 16;25(4):243-248. [doi: [10.1177/2010105816667137](https://doi.org/10.1177/2010105816667137)]
14. Roberts NJ, Kerr SM, Smith SM. Behavioral interventions associated with smoking cessation in the treatment of tobacco use. *Health Serv Insights* 2013 Aug 11;6:79-85 [FREE Full text] [doi: [10.4137/HSLI11092](https://doi.org/10.4137/HSLI11092)] [Medline: [25114563](https://pubmed.ncbi.nlm.nih.gov/25114563/)]
15. National Health Service. Motivation to Stop Smoking. NHS Center for Smoking Cessation and Training (NCSCCT). URL: <https://www.ncsct.co.uk/usr/pub/Motivation%20to%20stop%20smoking.pdf> [accessed 2020-10-05]
16. Iacobucci G. Stop smoking services: BMJ analysis shows how councils are stubbing them out. *BMJ* 2018 Aug 24;362:k3649 [FREE Full text] [doi: [10.1136/bmj.k3649](https://doi.org/10.1136/bmj.k3649)] [Medline: [30143515](https://pubmed.ncbi.nlm.nih.gov/30143515/)]
17. Filippidis F, Lavery A, Mons U, Jimenez-Ruiz C, Vardavas CI. Changes in smoking cessation assistance in the European Union between 2012 and 2017: pharmacotherapy versus counselling versus e-cigarettes. *Tob Control* 2019 Jan;28(1):95-100 [FREE Full text] [doi: [10.1136/tobaccocontrol-2017-054117](https://doi.org/10.1136/tobaccocontrol-2017-054117)] [Medline: [29563220](https://pubmed.ncbi.nlm.nih.gov/29563220/)]
18. Regmi K, Kassim N, Ahmad N, Tuah N. Effectiveness of Mobile Apps for Smoking Cessation: A Review. *Tob Prev Cessat* 2017 Apr 12;3(April):12 [FREE Full text] [doi: [10.18332/tpc/70088](https://doi.org/10.18332/tpc/70088)] [Medline: [32432186](https://pubmed.ncbi.nlm.nih.gov/32432186/)]
19. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2019 Oct 22;10:1-85 [FREE Full text] [doi: [10.1002/14651858.CD006611.pub5](https://doi.org/10.1002/14651858.CD006611.pub5)] [Medline: [31638271](https://pubmed.ncbi.nlm.nih.gov/31638271/)]
20. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. *J Med Internet Res* 2015 Jan 16;17(1):e17 [FREE Full text] [doi: [10.2196/jmir.3479](https://doi.org/10.2196/jmir.3479)] [Medline: [25596170](https://pubmed.ncbi.nlm.nih.gov/25596170/)]
21. Hoepfner BB, Hoepfner SS, Carlon HA, Perez GK, Helmuth E, Kahler CW, et al. Leveraging Positive Psychology to Support Smoking Cessation in Nondaily Smokers Using a Smartphone App: Feasibility and Acceptability Study. *JMIR Mhealth Uhealth* 2019 Jul 03;7(7):e13436 [FREE Full text] [doi: [10.2196/13436](https://doi.org/10.2196/13436)] [Medline: [31271147](https://pubmed.ncbi.nlm.nih.gov/31271147/)]

22. Tudor-Sfetea C, Rabee R, Najim M, Amin N, Chadha M, Jain M, et al. Evaluation of Two Mobile Health Apps in the Context of Smoking Cessation: Qualitative Study of Cognitive Behavioral Therapy (CBT) Versus Non-CBT-Based Digital Solutions. *JMIR Mhealth Uhealth* 2018 Apr 18;6(4):e98 [FREE Full text] [doi: [10.2196/mhealth.9405](https://doi.org/10.2196/mhealth.9405)] [Medline: [29669708](https://pubmed.ncbi.nlm.nih.gov/29669708/)]
23. Finkelstein J, Cha EM. Using a Mobile App to Promote Smoking Cessation in Hospitalized Patients. *JMIR Mhealth Uhealth* 2016 May 06;4(2):e59 [FREE Full text] [doi: [10.2196/mhealth.5149](https://doi.org/10.2196/mhealth.5149)] [Medline: [27154792](https://pubmed.ncbi.nlm.nih.gov/27154792/)]
24. Kappen D, Pejman M, Lennart N. Gamification of Older Adults' Physical Activity: An Eight-Week Study. 2018 Presented at: Proceedings of the 51st Hawaii International Conference on Systems Sciences; 2018; Hawaii p. 1207-1216 URL: <http://hdl.handle.net/10125/50036> [doi: [10.24251/HICSS.2018.149](https://doi.org/10.24251/HICSS.2018.149)]
25. Rajani NB, Weth D, Mastellos N, Filippidis FT. Adherence of popular smoking cessation mobile applications to evidence-based guidelines. *BMC Public Health* 2019 Jun 13;19(1):743 [FREE Full text] [doi: [10.1186/s12889-019-7084-7](https://doi.org/10.1186/s12889-019-7084-7)] [Medline: [31196062](https://pubmed.ncbi.nlm.nih.gov/31196062/)]
26. Sherwani Y. *Quit Genius* (2020). London, United Kingdom. URL: <https://www.quitgenius.com/> [accessed 2018-02-10]
27. Kretz G. *Kwit* (2020). Strasbourg, France. URL: <https://kwit.app/en> [accessed 2018-02-10]
28. Health statistics and information systems: Definition of regional groupings. World Health Organisation. URL: [https://www.who.int/healthinfo/global\\_burden\\_disease/definition\\_regions/en/](https://www.who.int/healthinfo/global_burden_disease/definition_regions/en/) [accessed 2020-09-10]
29. International Standard Classification of Education. International Labour Organisation (ILOSTAT). URL: <https://ilostat.ilo.org/resources/methods/classification-education/> [accessed 2020-08-20]
30. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict* 1991 Sep;86(9):1119-1127. [doi: [10.1111/j.1360-0443.1991.tb01879.x](https://doi.org/10.1111/j.1360-0443.1991.tb01879.x)] [Medline: [1932883](https://pubmed.ncbi.nlm.nih.gov/1932883/)]
31. van Amsterdam J, Vorspan F, Snijder M, van den Brink W, Schene AH, Stronks K, et al. Use of the Fagerström test to assess differences in the degree of nicotine dependence in smokers from five ethnic groups: The HELIUS study. *Drug Alcohol Depend* 2019 Jan 01;194:197-204 [FREE Full text] [doi: [10.1016/j.drugalcdep.2018.10.011](https://doi.org/10.1016/j.drugalcdep.2018.10.011)] [Medline: [30447512](https://pubmed.ncbi.nlm.nih.gov/30447512/)]
32. Etter J, Bergman M, Humair J, Perneger T. Development and validation of a scale measuring self-efficacy of current and former smokers. *Addiction* 2000 Jun;95(6):901-913. [doi: [10.1046/j.1360-0443.2000.9569017.x](https://doi.org/10.1046/j.1360-0443.2000.9569017.x)] [Medline: [10946439](https://pubmed.ncbi.nlm.nih.gov/10946439/)]
33. Motivation to Stop Smoking. NHS Centre for Smoking Cessation and Training. URL: <https://www.ncsc.co.uk/usr/pub/Motivation%20to%20stop%20smoking.pdf> [accessed 2018-02-18]
34. West R, Willis N. Double-blind placebo controlled trial of dextrose tablets and nicotine patch in smoking cessation. *Psychopharmacology (Berl)* 1998 Mar 11;136(2):201-204. [doi: [10.1007/s002130050557](https://doi.org/10.1007/s002130050557)] [Medline: [9551778](https://pubmed.ncbi.nlm.nih.gov/9551778/)]
35. Ussher M, Kakar G, Hajek P, West R. Dependence and motivation to stop smoking as predictors of success of a quit attempt among smokers seeking help to quit. *Addict Behav* 2016 Feb;53:175-180 [FREE Full text] [doi: [10.1016/j.addbeh.2015.10.020](https://doi.org/10.1016/j.addbeh.2015.10.020)] [Medline: [26547043](https://pubmed.ncbi.nlm.nih.gov/26547043/)]
36. Boardman T, Catley D, Mayo M, Ahluwalia J. Self-efficacy and motivation to quit during participation in a smoking cessation program. *Int J Behav Med* 2005;12(4):266-272 [FREE Full text] [doi: [10.1207/s15327558ijbm1204\\_7](https://doi.org/10.1207/s15327558ijbm1204_7)] [Medline: [16262545](https://pubmed.ncbi.nlm.nih.gov/16262545/)]
37. Cha S, Erar B, Niaura R, Graham A. Baseline Characteristics and Generalizability of Participants in an Internet Smoking Cessation Randomized Trial. *Ann Behav Med* 2016 Oct;50(5):751-761 [FREE Full text] [doi: [10.1007/s12160-016-9804-x](https://doi.org/10.1007/s12160-016-9804-x)] [Medline: [27283295](https://pubmed.ncbi.nlm.nih.gov/27283295/)]
38. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich A. Effective Behavioral Changes through a Digital mHealth App: Exploring the Impact of Hedonic Well-Being, Psychological Empowerment and Inspiration. *JMIR Mhealth Uhealth* 2018 Jun 15;6(6):e10024 [FREE Full text] [doi: [10.2196/10024](https://doi.org/10.2196/10024)] [Medline: [29907557](https://pubmed.ncbi.nlm.nih.gov/29907557/)]
39. Hiscock R, Bauld L, Amos A, Platt S. Smoking and socioeconomic status in England: the rise of the never smoker and the disadvantaged smoker. *J Public Health (Oxf)* 2012 Aug 28;34(3):390-396 [FREE Full text] [doi: [10.1093/pubmed/fds012](https://doi.org/10.1093/pubmed/fds012)] [Medline: [22375070](https://pubmed.ncbi.nlm.nih.gov/22375070/)]
40. Hiscock R, Dobbie F, Bauld L. Smoking Cessation and Socioeconomic Status: An Update of Existing Evidence from a National Evaluation of English Stop Smoking Services. *Biomed Res Int* 2015;2015:274056 [FREE Full text] [doi: [10.1155/2015/274056](https://doi.org/10.1155/2015/274056)] [Medline: [26273602](https://pubmed.ncbi.nlm.nih.gov/26273602/)]

---

## Abbreviations

**PID:** participant identification number

---

*Edited by L Buis; submitted 15.10.20; peer-reviewed by Y Lin, J Chan; comments to author 05.01.21; revised version received 18.01.21; accepted 02.04.21; published 30.04.21.*

*Please cite as:*

*Rajani NB, Mastellos N, Filippidis FT*

*Self-Efficacy and Motivation to Quit of Smokers Seeking to Quit: Quantitative Assessment of Smoking Cessation Mobile Apps*

*JMIR Mhealth Uhealth 2021;9(4):e25030*

*URL: <https://mhealth.jmir.org/2021/4/e25030>*

*doi: [10.2196/25030](https://doi.org/10.2196/25030)*

*PMID: [33929336](https://pubmed.ncbi.nlm.nih.gov/33929336/)*

©Nikita B Rajani, Nikolaos Mastellos, Filippos T Filippidis. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 30.04.2021. 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.

Original Paper

# Text Messaging Intervention for Young Smokers Experiencing Homelessness: Lessons Learned From a Randomized Controlled Trial

Sebastian Linnemayr<sup>1\*</sup>, PhD; Rushil Zutshi<sup>1\*</sup>, BTech, MS, MPhil; William Shadel<sup>2\*</sup>, PhD; Eric Pedersen<sup>3\*</sup>, PhD; Maria DeYoreo<sup>1\*</sup>, PhD; Joan Tucker<sup>1\*</sup>, PhD

<sup>1</sup>RAND Corporation, Santa Monica, CA, United States

<sup>2</sup>RAND Corporation, Pittsburgh, PA, United States

<sup>3</sup>University of Southern California, Los Angeles, CA, United States

\*all authors contributed equally

**Corresponding Author:**

Sebastian Linnemayr, PhD

RAND Corporation

1776 Main Street

Santa Monica, CA,

United States

Phone: 1 310 393 0411 ext 6734

Email: [slinnema@rand.org](mailto:slinnema@rand.org)

## Abstract

**Background:** Smoking rates are significantly higher among young people experiencing homelessness than in the general population. Despite a willingness to quit, homeless youth have little success in doing so on their own, and existing cessation resources tailored to this population are lacking. Homeless youth generally enjoy the camaraderie and peer support that group-based programs offer, but continuous in-person support during a quit attempt can be prohibitively expensive.

**Objective:** This study aimed to assess the feasibility and acceptability of an automated text messaging intervention (TMI) as an adjunct to group-based cessation counseling and provision of nicotine patches to help homeless youth quit smoking. This paper outlines the lessons learned from the implementation of the TMI intervention.

**Methods:** Homeless youth smokers aged 18 to 25 years who were interested in quitting (n=77) were recruited from drop-in centers serving homeless youth in the Los Angeles area. In this pilot randomized controlled trial, all participants received a group-based cessation counseling session and nicotine patches, with 52% (40/77) randomly assigned to receive 6 weeks of text messages to provide additional support for their quit attempt. Participants received text messages on their own phone rather than receiving a study-issued phone for the TMI. We analyzed baseline and follow-up survey data as well as back-end data from the messaging platform to gauge the acceptability and feasibility of the TMI among the 40 participants who received it.

**Results:** Participants had widespread (smart)phone ownership—16.4% (36/219) were ineligible for study participation because they did not have a phone that could receive text messages. Participants experienced interruptions in their phone use (eg, 44% [16/36] changed phone numbers during the follow-up period) but reported being able to receive the majority of messages. These survey results were corroborated by back-end data (from the program used to administer the TMI) showing a message delivery rate of about 95%. Participant feedback points to the importance of carefully crafting text messages, which led to high (typically above 70%) approval of most text messaging components of the intervention. Qualitative feedback indicated that participants enjoyed the group counseling session that preceded the TMI and suggested including more such group elements into the intervention.

**Conclusions:** The TMI was well accepted and feasible to support smoking cessation among homeless youth. Given high rates of smartphone ownership, the next generation of phone-based smoking cessation interventions for this population should consider using approaches beyond text messages and focus on finding ways to develop effective approaches to include group interaction using remote implementation. Given overall resource constraints and in particular the exigencies of the currently ongoing COVID-19 epidemic, phone-based interventions are a promising approach to support homeless youth, a population urgently in need of effective smoking cessation interventions.



**Trial Registration:** ClinicalTrials.gov NCT03874585; <https://clinicaltrials.gov/ct2/show/NCT03874585>

**International Registered Report Identifier (IRRID):** RR2-10.1186/s13722-020-00187-6

(*JMIR Mhealth Uhealth* 2021;9(4):e23989) doi:[10.2196/23989](https://doi.org/10.2196/23989)

## KEYWORDS

smokers; quitting; text messaging; homeless; young adults; cessation resources; peer support; smoking rate; smartphone

## Introduction

National data indicate that 19% of people aged 18 to 25 years in the United States are current (past 30 day) cigarette smokers [1]. However, rates of smoking among young people experiencing homelessness are significantly higher, with several studies indicating that up to 70% of the population are current smokers [2-4]. They also spend a higher fraction (about 30%) [5] of their monthly income on cigarettes, compared with about 20% spent by homeless adult smokers [6]. In addition, most homeless youth who smoke report engaging in one or more high-risk smoking practices that heighten their exposure to toxins such as smoking shared cigarettes (96%), smoking discarded butts (71%) and filters (46%), and blocking filter vents (39%) [7]. Despite the high prevalence of smoking and high-risk smoking behaviors negatively impacting the health of homeless youth who smoke, evidence-based smoking cessation programs specifically designed for this population are lacking.

This paper is based on our previous research that focused specifically on young homeless smokers and supported the need for cessation programs adapted to this population. We found high willingness to quit smoking among this population but little success in doing so on their own or when using existing available cessation resources [5,8-10]. In a sample of nearly 300 young homeless smokers, almost half (43%) were motivated to quit in the next 30 days and, of those, 76% were interested in using a nicotine replacement product (eg, nicotine patches) and/or smoking cessation counseling to help them quit. The participants reported attempting to quit an average of nearly 10 times in the past year, and two-thirds of them had quit for at least 24 hours before relapsing. Most participants (79%) who had tried to quit smoking did so on their own (ie, without counseling or medication), reflecting the lack of readily available smoking cessation services for this population. As many of these young people are not connected to the formal health care sector, we identified drop-in centers as a low barrier, “come as you are” point of service entry for young homeless smokers. These centers represent an ideal setting to reach young homeless smokers who may not seek services elsewhere, and we therefore recruited the sample for this study in this setting.

We also found that homeless youth who are interested in quitting enjoy the camaraderie and peer support that group-based programs offer [8]. Therefore, we decided to use a group format to deliver a brief in-person smoking cessation counseling session at the beginning of the intervention as described below. While there is no clear consensus on the cost effectiveness of mobile health (mHealth) behavioral interventions [11,12], in this case, continued in-person contact during a quit attempt is prohibitively time- and cost-intensive, and instead support must be provided

in a way that is low cost and low burden for both service providers and their clients. A text messaging-based intervention (TMI) for cessation support circumvents some of these barriers and offers a tailored approach that can be accessed anywhere and anytime. Given widespread cell phone ownership among homeless youth, phone-based support holds great potential for behavior change [13]. Although TMIs for behavior change have not been evaluated in this population, a recent study using text messaging for daily data collection among young homeless smokers found it to be both acceptable and feasible [14]. Systematic reviews of text messaging services for adolescents have shown the general acceptability and feasibility as well as improvements in preventive behaviors and adherence to medication in the case of chronic health conditions [15,16]. Furthermore, nearly all of the homeless youth in our previous study expressed interest in using their phone to receive ongoing text messages to help them quit [10]. Phone-based smoking cessation interventions have been found to be feasible and effective in other populations, as a recent meta-analysis of 22 studies found that smokers who received a TMI were more likely to abstain from smoking relative to controls across a number of outcomes, including 7-day point prevalence (odds ratio [OR] 1.38, 95% CI 1.22-1.55) and continuous abstinence (OR 1.63, 95% CI 1.19-2.24) [17]. TMIs hold great potential (even more so in the age of COVID-19) for offering ongoing support for behavior change as an adjunct to face-to-face services.

In this paper, we report on our experiences implementing the first TMI targeted at smoking cessation among homeless youth as a part of a pilot evaluation that found promising results of the TMI on smoking reduction. We lay out the lessons learned in designing appropriate and effective text messages and implementing the technical aspects of the intervention and report on the feedback from the participants regarding different aspects of the intervention. We conclude with a discussion of what worked well and what could benefit from improvement with the hope that our lessons can be of value to other, much-needed TMIs for this particularly vulnerable population.

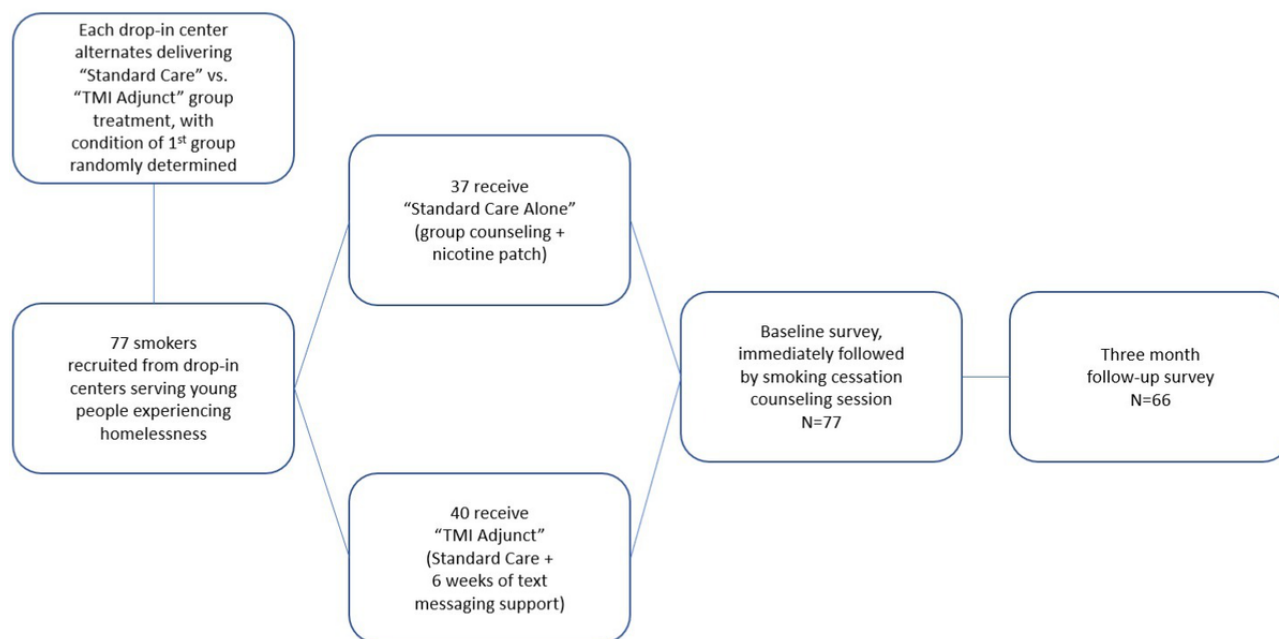
## Methods

### Study Design and Participants

The findings reported here are from a pilot study of 77 current smokers aged 18 to 25 years who desired to quit and were recruited from 3 drop-in centers serving homeless youth in the Los Angeles area. The unit of analysis was the individual, but individuals were assigned to groups (standard care alone vs TMI adjunct) based on the drop-in center where they were seeking services during recruitment hours. The intervention was carried out in a cluster cross-over randomized controlled design

such that each drop-in center alternated between serving as an intervention or control site by phase across the field period. After formative work with 26 participants [18], we settled on a study design that provided a brief group counseling session based on the 5 As (ask, advise, assess, assist, and arrange) [19] and nicotine patches to all participants, and those in the TMI condition were then sent automated text messages. The study flow is shown in Figure 1, and a detailed description of the intervention design has been published elsewhere [20]; here we focus on aspects related to the text messaging component.

**Figure 1.** Study flow.



## Description of TMI

The project team generated 174 text messages to send to all TMI participants (see previous work [20] for example texts by the 5 main foci of the TMI program as well as TMI time point [ie, pre-quit day texts, quit day texts, early post-quit texts, later post-quit texts]). Text message content was informed by our prior work with homeless smokers to identify key factors associated with motivation to quit, consultation with the text messaging literature, and review of text messages included in other public domain smoking cessation programs (eg, Text2Quit [21] and SmokefreeTXT [22]). The majority of texts address 1 of the 5 main foci of the intervention that are based on factors identified in our prior work associated with motivation to quit among young homeless smokers [5]: strategies for getting support for quitting ("Check in with your friends and let them know how you're doing with staying smoke-free. If you need support from them, ask for it!"); calculations for the amount of money saved by quitting ("Quitting smoking can save you some big \$\$! If you're curious about how much you can save by quitting, check out this link: [link to external website]"); presentation of health and social benefits of quitting ("Sniping discarded butts might be free tobacco, but it can also make you sick from germs on the ground or from the person who smoked it first. Not worth it."); strategies for dealing with cravings and

Enrollment and intervention delivery occurred between May 30, 2019, and February 25, 2020, with follow-up data collection occurring approximately 90 days later. The sample was 81% male (62/77), 71% heterosexual (55/77), and 13% non-Hispanic White (10/77). Mean age was 22.61 (SD 1.84) years. All study materials and procedures were approved by the study's institutional review board, and a certificate of confidentiality was obtained from the National Institutes of Health. The study was registered at ClinicalTrials.gov [NCT03874585].

negative moods ("Don't lose the progress you've made. Ride through cravings by chewing gum, walking it off, or taking 10 deep breaths. And text CRAVE for more support anytime"); and tips for staying motivated ("Say out loud 'I'm a nonsmoker.' It seems cheesy, but it can remind you of all the changes you've made and help you stay strong through the cravings"). Given the provision of nicotine patches in this study, some texts included reminders to use the nicotine patches. Other text messages included periodic check-ins to see if they were still reading the texts ("We just want to know that you got this text. Please text back: YES") and occasional fun content (eg, encouraging messages, emojis, or funny memes). Similar to other text messaging programs (eg, SmokefreeTXT), participants could text CRAVE, MOOD, or SLIP at any time for additional support in dealing with cravings, negative mood, or a smoking lapse, respectively; an additional 72 text messages (24 for each text type) were developed for these purposes.

It was important to ensure that the content of the text messages reflected the unique circumstances of young people experiencing homelessness. In addition, to make messages more effective we used recent insights from behavioral economics in the design of the messages [23,24], such as using gain/loss framing ("Increase your chances of success by starting to use the nicotine patches tomorrow morning"), employing social norms ("You're

not alone in wanting to be smoke-free. Most people your age don't smoke...and most people that do smoke want to quit."), appealing to participants' self-identity ("Look in the mirror and tell yourself: I am a nonsmoker! Thinking of yourself as a 'nonsmoker' instead of an 'ex-smoker' can make staying smoke-free easier."), and providing progress information to keep the salience of the quitting behavior high ("You've worked really hard for almost 2 weeks to get where you are now. Don't lose this energy!"). The program was named CRUSH IT! in this spirit as well.

Given that TMIs have not been previously implemented for behavior change among young homeless smokers, there is little guidance from the existing literature regarding the frequency and timing of text messages for this population. Therefore, we conducted several focus groups with a total of 18 homeless smokers and elicited usability testing feedback with a separate sample of 10 homeless smokers recruited from the drop-in centers [18]. During these groups we also reviewed the content of texts with participants and made sure that receiving texts would be feasible for this population to inform decisions about the optimal content (eg, what types of messages would likely be most effective in dealing with cravings) and wording (eg, how to word the text message so that it motivates homeless youth to stay quit). During conceptualization of the TMI, we decided to use the online text messaging platform Telerivet to deliver the intervention. Telerivet is a free online web service that one of the authors used in previous studies [25-27]. Telerivet also offers a paid version with additional features that we made use of in our intervention. For instance, the paid version allows users to create multiple (up to 20) polls. These polls are useful for asking questions to the participants and setting up automated responses while also logging the participants' responses. We had polls set up to occasionally check in with the participants to see if they had smoked that day and to respond with words of appreciation or encouragement based on their responses. Early on, we also used these polls to log their reason for wanting to quit smoking and sent it to them later on during the intervention to remind them of what they had said. The paid version of the platform also allows for a larger degree of automation and backend data to be collected. Since the TMI had a different start date for each intervention group, reprogramming the platform with all the messages every time would have been prohibitively labor intensive. Instead, we used Telerivet features to completely automate this process using triaged scheduling relative to the start date of each group. Telerivet also allows for a great degree of customizability to create functionalities beyond the preprogrammed options. For example, we programmed a feature allowing us to create keywords that users could send in order to get automatic supportive texts if they were tempted to smoke due to a craving or negative mood or if they had slipped and smoked a cigarette. We were also able to automate the process for allowing participants to delay the start of the intervention by a few days or stop it altogether.

## Surveys

The results in the following section come from two surveys. The first was a baseline survey that was administered when participants were first recruited for the TMI before the counseling session. This survey asked respondents for information such as demographics, phone use, smoking behavior at baseline, and other substance use. Three months after the baseline survey (and hence about 6 weeks after the end of the intervention phase), all participants received a follow-up survey that asked for information similar to that at baseline; those in the TMI condition were asked additional questions on their experience with the TMI. The baseline survey was administered in a group setting immediately prior to the group counseling session. We used self-administered paper-pencil forms, which was the most feasible option in our field setting. The follow-up surveys were administered either in person or via phone interview. Survey response forms were then scanned and checked for accuracy. While staff were available to assist participants in completing the surveys, no one required such assistance. Participants received \$20 for the baseline survey and \$40 for the follow-up survey. The baseline survey had 77 participants while the follow-up survey had 66 participants due to attrition.

## Results

### Lesson 1. Young Homeless Smokers Had Widespread (Smart)phone Ownership

In [Table 1](#), we present the descriptive statistics regarding phone ownership and other phone-related parameters of participants in the standard and TMI conditions. Randomization appeared successful as there were no statistically significant differences between the two groups; although given the pilot nature of this study, we were not powered to find statistically significant differences. A key eligibility criterion was having a phone with them at the recruitment visit that could receive text messages. Confirming our assertion that we expected widespread phone ownership, of the youth who were asked about having a phone, only 16.4% (36/219) reported that they did not have one with them that could receive text messages. In addition, 3 who had a phone reported that they did not want to receive text messages (for a total of 39/219 (17.8%). Smartphone ownership was high—84% (31/37) in the standard condition and 73% (29/40) in the TMI condition owned smartphones. A majority of participants—62% (23/37) in the standard condition and 78% (31/40) in the TMI condition—had unlimited minutes. More importantly for our TMI, a majority also had unlimited texts—68% (25/37) in the standard condition and 80% (32/40) in the TMI condition. About half of the participants—40% (15/37) in the standard condition and 55% (22/40) in the TMI condition—had unlimited data. Many had obtained their phone through a benefits program such as the Lifeline program, which gives low-income Americans free cell phones, voice minutes, and texting—32% (13/37) in the standard condition and 52% (21/40) in the TMI condition.

**Table 1.** Participant cellphone characteristics at baseline.

Variables	Standard condition (answered yes; n=37), n (%)	TMI <sup>a</sup> condition (answered yes; n=40), n (%)
It is my own phone	32 (86)	34 (85)
It is a phone owned by a friend, partner, or relative	5 (13)	6 (15)
It is a phone obtained through a benefits program	13 (32)	21 (52)
It is a smartphone	31 (84)	29 (73)
I have unlimited minutes on this phone	23 (62)	31 (78)
I have unlimited texts on this phone	25 (68)	32 (80)
I have unlimited data on this phone	15 (40)	22 (55)
I can view webpages on this phone because I have a data plan	27 (73)	27 (67)
I can only view webpages on this phone when I'm connected to the Wi-Fi	15 (40)	20 (50)

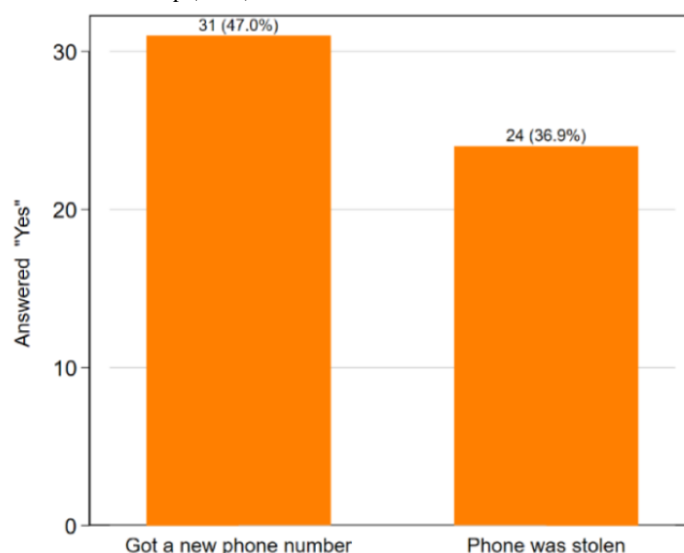
<sup>a</sup>TMI: text messaging-based intervention.

## Lesson 2. Young Homeless Smokers Experienced Interruptions in Their Phone Use But Were Able to Receive the Majority of Messages

As can be seen in [Figure 2](#), there were phone-related challenges that could influence the effectiveness of a TMI smoking cessation program. Almost half (31/66, 47%) of the sample changed phone numbers in the approximately 3 months between recruitment and the follow-up survey (15/30 [50%] in the standard condition and 16/36 [44%] in the TMI condition). Some (24/66, 37%) participants reported having their phone stolen during this period (11/30 [37%] in the standard condition and 13/36 [36%] in the TMI condition). However, based on feedback from the interviewers, a majority of participants who reported getting a new number did so between the end of the 6-week TMI period and the follow-up survey in month 3 (ie, outside of the intervention period). In addition, for participants in the TMI condition, a majority (26/36, 72%) had trouble charging their phones at least sometimes, and 33% (12/36) had no reception at some point during that 3-month period.

Despite the frequent occurrence of phone loss and switching of numbers, 81% (29/36) of TMI participants reported not having

problems receiving the texts. Further, 72% (26/36) had no trouble accessing the hyperlinks provided in some of the text messages, allowing them to access supplemental information (see [Table 2](#)). Moreover, backend Telerivet data indicated a 95% message sending success rate (ie, about 95% of messages were successfully delivered to the participants on average). The reason for nondelivery of messages was either because participants' phones were switched off for an extended period or potentially due to service issues with the carrier. A total of 55% (20/36) of TMI participants responded to at least one prompt over the 6-week period, and 11% (4/36) texted one of the unprompted keywords such as MOOD or CRAVE. Three participants proactively reached out to us to report having gotten a new phone number (participants were told at enrollment that they would receive a \$5 incentive for providing updated contact information); others who got new numbers were successfully contacted using social media or other means for purposes of follow-up survey data collection. Consequently, the sample sizes between [Table 1](#) and [Figure 2](#) differ since follow-up data was collected for 30 standard condition and 36 TMI condition participants.

**Figure 2.** Cell phone and number retention at follow-up (n=66).

**Table 2.** Intervention delivery metrics.

Metric	Never, n (%)	Sometimes, n (%)	Often, n (%)	Always, n (%)
Trouble keeping phone battery charged	10 (28)	17 (47)	6 (17)	3 (8)
Ran out of messages and couldn't get messages	29 (80)	3 (8)	2 (6)	2 (6)
Trouble accessing hyperlinks on phone	26 (72)	4 (11)	2 (6)	4 (11)
Did not have cellphone reception	24 (66)	10 (28)	1 (3)	1 (3)

### Lesson 3. Carefully Crafting Text Messages Led to High Approval of Most Text Messaging Components Among the Participants

Table 3 presents the opinions of participants about different TMI components. Most participants somewhat or strongly agreed that they would recommend the program to a friend interested in quitting (30/36, 83%); liked being able to text CRAVE, MOOD, or SLIP at any time to get additional support (26/36, 72%); and found that the messages helped when they were experiencing a craving (25/36, 69%). Taken together, these findings indicate participants felt that the intervention provided appropriate support for alleviating or navigating the temptation to smoke. Nearly three-quarters (26/36, 72%) of participants found the text messages helpful and said the hyperlinks provided them with useful information. Further, it appears that customizing the intervention to cater to this population was also

well received—66% (24/36) reported liking the tone of the message, 78% (28/36) liked that the messages were personalized (such as by including names or referring to previous responses), and 75% (27/36) liked the use of memes and emojis in the messages. A significant proportion of participants (14/36, 44%) reported receiving the messages late because their phone was switched off or had bad reception. Further, we had open-ended questions to allow for more detailed feedback on what they liked most and least about CRUSH IT! Participants mentioned appreciating the attention to detail, the resources provided in the text messages, and the positive tone of the messages. In terms of areas for improvement, the most consistent feedback was wanting more of a group component to the ongoing support, which was mentioned by several participants. In addition, 2 participants remarked on the timing of the messages, one noted the lack of a real person for support and questions, and one thought that too many messages were sent.

**Table 3.** Intervention acceptability metrics.

Metric	Strongly disagree, n (%)	Somewhat disagree, n (%)	Somewhat agree, n (%)	Strongly agree, n (%)
<b>The text messages:</b>				
helped when I was experiencing a craving	4 (11)	6 (17)	17 (49)	8 (23)
helped when I was having trouble staying quit	3 (9)	9 (27)	14 (41)	8 (23)
helped keep me motivated to quit	3 (9)	6 (17)	11 (31)	15 (43)
provided me information that I could use	6 (17)	3 (9)	4 (11)	22 (63)
<b>I liked:</b>				
being able to text a keyword for extra support	5 (14)	4 (11)	7 (20)	19 (53)
the memes and emojis that were sent	4 (11)	4 (11)	10 (29)	17 (49)
that the text messages were personalized	4 (12)	2 (6)	11 (32)	17 (50)
the tone of the text messages	6 (18)	4 (12)	12 (35)	12 (35)
when the messages asked me to respond	6 (18)	5 (14)	14 (41)	9 (27)
I received text messages late because my phone wasn't charged, or I had bad reception	14 (41)	5 (14)	9 (27)	5 (14)
Hyperlinks to websites provided information I could use	5 (14)	3 (9)	10 (29)	16 (47)
I would recommend the CRUSH IT! program to a friend who is trying to quit smoking	4 (12)	1 (3)	7 (19)	23 (64)

## Discussion

### Principal Findings

In this paper, we discussed the development of a smoking cessation TMI for young homeless smokers and reported lessons learned regarding intervention components such as development of the messages, (smart)phone ownership, technical challenges

with the intervention, and feasibility (such as whether messages were received) and acceptability (ie, feedback about what intervention aspects the participants liked).

Regarding message development, we found that conducting focus groups with young homeless smokers and pilot testing with participants before rolling out the intervention were crucial steps to successfully tailoring the content, tone, and frequency

of messages to the target population. Participants preferred messages that provided information specific to their current living situation, were light in tone rather than being preachy, and contained fun elements such as emojis and memes. We then programmed the messages developed with the participants' input using a web-based text messaging platform that allowed for automatized sending of messages and categorization of participant responses for subsequent sending of appropriate follow-up messages, which minimized human resources required for implementation and guaranteed that messages were sent out with a high success rate.

As would be expected based on other research showing widespread cell phone ownership among individuals experiencing homelessness [28-30], we found that we had to exclude only a small percentage of young people experiencing homelessness from participating due to not having a phone that could receive text messages; indeed, most participants had smartphones, unlimited text and call plans, and a significant portion even had unlimited data plans. Taken together, these results suggest that phone and even smartphone-based interventions are feasible for this highly mobile population. Reflecting the challenges of cell phone ownership among individuals experiencing homelessness reported previously [31], during the intervention we realized that a large fraction of participants either had their phone stolen or switched numbers and most had trouble keeping their phone consistently charged. However, based on participant feedback they still received a majority of text messages, speaking to their resilience in the face of technical difficulties but also to the relatively short intervention duration of 6 weeks during which most participants may not have yet experienced these problems (the data were collected several weeks after the intervention ended and may have picked up difficulties arising after the intervention period).

Last, similar to prior research finding high acceptability among young people experiencing homelessness for using text messaging for daily data collection [14], we found high acceptability of our text message support for quitting smoking, with a majority of participants indicating that they would recommend the TMI to a friend trying to quit. Participants found the messages to contain information useful to them and they liked the light tone we tried to convey and the fun elements such as emojis and memes, demonstrating that the formative work and piloting of the intervention we performed before rolling it out to all participants were important to hit the right content and tone. Given increasing evidence of technology fatigue and the failure of many individuals to either take up or keep engaged with mHealth interventions [32], this is an important lesson for future such interventions. Overall, our results show that text message-based interventions can be a valuable approach to support smoking cessation efforts of homeless youth who are highly mobile and difficult (and costly) to reach with traditional means, and that the specific needs and preferences of this population need to be taken into account in the design and implementation of such an intervention.

## Strengths and Limitations

Our study has significant strengths such as being the first TMI for smoking cessation support for homeless youth and implementation at 3 sites in the Los Angeles area, which allowed for testing in different contexts. However, it also had limitations. For example, results may not generalize to homeless youth in other geographic areas or to those younger than age 18 years. Due to the study being a pilot, we were intentionally not powered to evaluate statistical differences between the standard and TMI conditions. Future studies should test the promising results of this pilot in a fully powered trial ideally in several different locations to test for generalizability of our results. In addition, it is a limitation that we did not collect qualitative data (eg, debriefing interviews) with the homeless youth who used the TMI at the end of the study, which might have identified additional strengths and weaknesses of this approach that could inform future research efforts in this area. Clearly, our study contributes to a growing evidence base that TMIs for smoking cessation can be effectively implemented even for highly transient and resource-constrained populations such as homeless youth, but that adaptation to their specific needs (including conducting appropriate formative work) is needed to render the TMI acceptable.

In terms of future directions, one specific recommendation we got from participant feedback is that they really enjoyed the group cessation counseling session prior to receiving the TMI and wanted more group elements incorporated into the intervention. Going forward, future smoking cessation interventions for this population should consider using approaches making full use of smartphone capabilities, including virtual approaches to leverage group interaction not requiring in-person meetings such as using social media chatrooms. Given their low costs and low requirements for human resources (particularly given the currently ongoing COVID-19 epidemic), (smart)phone-based interventions are a promising approach to support homeless youth, a population urgently in need of effective smoking cessation interventions.

## Conclusions

In conclusion, we find that most young homeless smokers have cell phones that allow mHealth interventions, with many being in possession of smartphones that typically have unlimited minutes and texts. Given that almost half also have unlimited data plans, it seems that in the near future internet-based interventions requiring smartphones will also be feasible in this population. In line with widespread concern in the literature, we find that homeless youth reported frequent occurrence of phone loss and switching of numbers. However, despite these difficulties the majority of participants reported not having problems receiving the study texts, and engagement with the different intervention components was generally high. We hope that the lessons derived from this pilot intervention serve as useful inputs for future mHealth studies for this population in need of smoking cessation interventions.

## Acknowledgments

This research was supported by funds from Tobacco-Related Disease Research Program of the University of California (grant number 27IP-0051; principal investigator JT). The authors wish to thank the drop-in centers that allowed us to conduct this research, Isabel Leamon, Alice Kim, and Sarika Bharil for their assistance with enrollment and program delivery, and Rick Garvey and the RAND Survey Research Group for their assistance in conducting the follow-up data collection. Our thanks to the study participants who so generously volunteered their time and insight to allow this study to happen.

## Conflicts of Interest

None declared.

## References

1. Key substance use and mental health indicators in the United States: results from the 2017 National Survey on Drug Use and Health. Rockville: Substance Abuse and Mental Health Services Administration; 2017. URL: <https://store.samhsa.gov/sites/default/files/d7/priv/sma18-5068.pdf> [accessed 2021-02-26]
2. Wenzel SL, Tucker JS, Golinelli D, Green HD, Zhou A. Personal network correlates of alcohol, cigarette, and marijuana use among homeless youth. *Drug Alcohol Depend* 2010 Nov 01;112(1-2):140-149 [FREE Full text] [doi: [10.1016/j.drugalcdep.2010.06.004](https://doi.org/10.1016/j.drugalcdep.2010.06.004)] [Medline: [20656423](https://pubmed.ncbi.nlm.nih.gov/20656423/)]
3. Baer JS, Ginzler JA, Peterson PL. DSM-IV alcohol and substance abuse and dependence in homeless youth. *J Stud Alcohol* 2003 Jan;64(1):5-14. [doi: [10.15288/jsa.2003.64.5](https://doi.org/10.15288/jsa.2003.64.5)] [Medline: [12608478](https://pubmed.ncbi.nlm.nih.gov/12608478/)]
4. Bousman CA, Blumberg EJ, Shillington AM, Hovell MF, Ji M, Lehman S, et al. Predictors of substance use among homeless youth in San Diego. *Addict Behav* 2005 Jul;30(6):1100-1110. [doi: [10.1016/j.addbeh.2004.10.006](https://doi.org/10.1016/j.addbeh.2004.10.006)] [Medline: [15925120](https://pubmed.ncbi.nlm.nih.gov/15925120/)]
5. Tucker JS, Shadel WG, Golinelli D, Ewing B, Mullins L. Motivation to quit and interest in cessation treatment among homeless youth smokers. *Nicotine Tob Res* 2015 Aug;17(8):990-995. [doi: [10.1093/ntr/ntu261](https://doi.org/10.1093/ntr/ntu261)] [Medline: [26180224](https://pubmed.ncbi.nlm.nih.gov/26180224/)]
6. Arnsten JH, Reid K, Bierer M, Rigotti N. Smoking behavior and interest in quitting among homeless smokers. *Addict Behav* 2004 Aug;29(6):1155-1161. [doi: [10.1016/j.addbeh.2004.03.010](https://doi.org/10.1016/j.addbeh.2004.03.010)] [Medline: [15236817](https://pubmed.ncbi.nlm.nih.gov/15236817/)]
7. Tucker JS, Shadel WG, Golinelli D, Mullins L, Ewing B. Sniping and other high-risk smoking practices among homeless youth. *Drug Alcohol Depend* 2015 Sep 01;154:105-110. [doi: [10.1016/j.drugalcdep.2015.06.036](https://doi.org/10.1016/j.drugalcdep.2015.06.036)] [Medline: [26160458](https://pubmed.ncbi.nlm.nih.gov/26160458/)]
8. Mullins L, O'Hanlon CE, Shadel WG, Tucker JS. A qualitative study of smoking cessation experiences and perceptions among homeless young adults. *J Soc Distress Homeless* 2017 Sep 14;27(1):1-8. [doi: [10.1080/10530789.2017.1377959](https://doi.org/10.1080/10530789.2017.1377959)]
9. Shadel WG, Tucker JS, Mullins L, Staplefoote L. Providing smoking cessation programs to homeless youth: the perspective of service providers. *J Subst Abuse Treat* 2014 Oct;47(4):251-257. [doi: [10.1016/j.jsat.2014.05.009](https://doi.org/10.1016/j.jsat.2014.05.009)] [Medline: [25012554](https://pubmed.ncbi.nlm.nih.gov/25012554/)]
10. Tucker JS, Shadel WG, Golinelli D, Ewing B, Mullins L, Staplefoote BL. Reducing cigarette smoking among unaccompanied homeless youth. Rand Corporation. 2015. URL: [https://www.rand.org/content/dam/rand/pubs/research\\_briefs/RB9800/RB9828/RAND\\_RB9828.pdf](https://www.rand.org/content/dam/rand/pubs/research_briefs/RB9800/RB9828/RAND_RB9828.pdf) [accessed 2021-02-26]
11. 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](https://doi.org/10.2196/mhealth.6425)] [Medline: [27780795](https://pubmed.ncbi.nlm.nih.gov/27780795/)]
12. 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](https://doi.org/10.1371/journal.pone.0170581)] [Medline: [28152012](https://pubmed.ncbi.nlm.nih.gov/28152012/)]
13. Rice E, Lee A, Taitt S. Cell phone use among homeless youth: potential for new health interventions and research. *J Urban Health* 2011 Dec;88(6):1175-1182 [FREE Full text] [doi: [10.1007/s11524-011-9624-z](https://doi.org/10.1007/s11524-011-9624-z)] [Medline: [22076445](https://pubmed.ncbi.nlm.nih.gov/22076445/)]
14. Tyler KA, Schmitz RM. Using cell phones for data collection: benefits, outcomes, and intervention possibilities with homeless youth. *Child Youth Serv Rev* 2017 Mar;76:59-64 [FREE Full text] [doi: [10.1016/j.childyouth.2017.02.031](https://doi.org/10.1016/j.childyouth.2017.02.031)] [Medline: [28970644](https://pubmed.ncbi.nlm.nih.gov/28970644/)]
15. 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](https://doi.org/10.2196/mhealth.7798)] [Medline: [28506955](https://pubmed.ncbi.nlm.nih.gov/28506955/)]
16. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: a systematic review. *JMIR Mhealth Uhealth* 2017 Apr 19;5(4):e50 [FREE Full text] [doi: [10.2196/mhealth.6837](https://doi.org/10.2196/mhealth.6837)] [Medline: [28428157](https://pubmed.ncbi.nlm.nih.gov/28428157/)]
17. Scott-Sheldon LAJ, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2016;4(2):e49 [FREE Full text] [doi: [10.2196/mhealth.5436](https://doi.org/10.2196/mhealth.5436)] [Medline: [27207211](https://pubmed.ncbi.nlm.nih.gov/27207211/)]
18. Tucker JS, Pedersen ER, Linnemayr S, Shadel WG, DeYoreo M, Zutshi R. A text message intervention for quitting cigarette smoking among young adults experiencing homelessness: study protocol for a pilot randomized controlled trial. *Addict Sci Clin Pract* 2020 Feb 19;15(1):11 [FREE Full text] [doi: [10.1186/s13722-020-00187-6](https://doi.org/10.1186/s13722-020-00187-6)] [Medline: [32075695](https://pubmed.ncbi.nlm.nih.gov/32075695/)]

19. Shadel WG, Niaura R. Brief behavioral interventions. In: Abrams DB, Niaura R, Brown R, Emmons K, Goldstein MG, editors. *The Tobacco Dependence Treatment Handbook: A Guide to Best Practices*. New York: Guilford Press; 2003:101-117.
20. Tucker JS, Pedersen ER, Linnemayr S, Shadel WG, DeYoreo M, Zutshi R. A text message intervention for quitting cigarette smoking among young adults experiencing homelessness: study protocol for a pilot randomized controlled trial. *Addict Sci Clin Pract* 2020 Feb 19;15(1):1-13 [[FREE Full text](#)] [doi: [10.1186/s13722-020-00187-6](https://doi.org/10.1186/s13722-020-00187-6)] [Medline: [32075695](https://pubmed.ncbi.nlm.nih.gov/32075695/)]
21. Abrams LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of Text2Quit: a text messaging program for smoking cessation. *Am J Prev Med* 2014 Sep;47(3):242-250. [doi: [10.1016/j.amepre.2014.04.010](https://doi.org/10.1016/j.amepre.2014.04.010)] [Medline: [24913220](https://pubmed.ncbi.nlm.nih.gov/24913220/)]
22. SmokefreeTXT. URL: <https://smokefree.gov/smokefreetxt> [accessed 2021-03-22]
23. Garcia-Retamero R, Cokely ET. Effective communication of risks to young adults: using message framing and visual aids to increase condom use and STD screening. *J Exp Psychol Appl* 2011 Sep;17(3):270-287. [doi: [10.1037/a0023677](https://doi.org/10.1037/a0023677)] [Medline: [21942316](https://pubmed.ncbi.nlm.nih.gov/21942316/)]
24. Linnemayr S, O'Hanlon C, Uscher-Pines L, Van Abel K, Nelson C. Using insights from behavioral economics to strengthen disaster preparedness and response. *Disaster Med Public Health Prep* 2016 Oct;10(5):768-774. [doi: [10.1017/dmp.2016.29](https://doi.org/10.1017/dmp.2016.29)] [Medline: [27188246](https://pubmed.ncbi.nlm.nih.gov/27188246/)]
25. MacCarthy S, Mendoza-Graf A, Huang H, Mukasa B, Linnemayr S. Supporting Adolescents to Adhere (SATA): lessons learned from an intervention to achieve medication adherence targets among youth living with HIV in Uganda. *Child Youth Serv Rev* 2019 Jul;102:56-62 [[FREE Full text](#)] [doi: [10.1016/j.childyouth.2019.04.007](https://doi.org/10.1016/j.childyouth.2019.04.007)] [Medline: [31223179](https://pubmed.ncbi.nlm.nih.gov/31223179/)]
26. Linnemayr S, Huang H, Luoto J, Kambugu A, Thirumurthy H, Haberer JE, et al. Text messaging for improving antiretroviral therapy adherence: no effects after 1 year in a randomized controlled trial among adolescents and young adults. *Am J Public Health* 2017 Dec;107(12):1944-1950. [doi: [10.2105/AJPH.2017.304089](https://doi.org/10.2105/AJPH.2017.304089)] [Medline: [29048966](https://pubmed.ncbi.nlm.nih.gov/29048966/)]
27. Rana Y, Haberer J, Huang H, Kambugu A, Mukasa B, Thirumurthy H, et al. Short message service (SMS)-based intervention to improve treatment adherence among HIV-positive youth in Uganda: focus group findings. *PLoS One* 2015;10(4):e0125187 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0125187](https://doi.org/10.1371/journal.pone.0125187)] [Medline: [25881059](https://pubmed.ncbi.nlm.nih.gov/25881059/)]
28. Jennings L, Lee N, Shore D, Strohminger N, Allison B, Conserve DF, et al. U.S. minority homeless youth's access to and use of mobile phones: implications for mhealth intervention design. *J Health Commun* 2016 Jul;21(7):725-733. [doi: [10.1080/10810730.2015.1103331](https://doi.org/10.1080/10810730.2015.1103331)] [Medline: [27232544](https://pubmed.ncbi.nlm.nih.gov/27232544/)]
29. Rhoades H, Wenzel S, Rice E, Winetrobe H, Henwood B. No digital divide? Technology use among homeless adults. *J Soc Distress Homeless* 2017;26(1):73-77 [[FREE Full text](#)] [doi: [10.1080/10530789.2017.1305140](https://doi.org/10.1080/10530789.2017.1305140)] [Medline: [31097900](https://pubmed.ncbi.nlm.nih.gov/31097900/)]
30. Post LA, Vaca FE, Doran KM, Luco C, Naftilan M, Dziura J, et al. New media use by patients who are homeless: the potential of mHealth to build connectivity. *J Med Internet Res* 2013 Sep 03;15(9):e195 [[FREE Full text](#)] [doi: [10.2196/jmir.2724](https://doi.org/10.2196/jmir.2724)] [Medline: [24001876](https://pubmed.ncbi.nlm.nih.gov/24001876/)]
31. Raven MC, Kaplan LM, Rosenberg M, Tieu L, Guzman D, Kushel M. Mobile phone, computer, and internet use among older homeless adults: results from the HOPE HOME cohort study. *JMIR Mhealth Uhealth* 2018 Dec 10;6(12):e10049 [[FREE Full text](#)] [doi: [10.2196/10049](https://doi.org/10.2196/10049)] [Medline: [30530464](https://pubmed.ncbi.nlm.nih.gov/30530464/)]
32. 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;22(5):e17572 [[FREE Full text](#)] [doi: [10.2196/17572](https://doi.org/10.2196/17572)] [Medline: [32348255](https://pubmed.ncbi.nlm.nih.gov/32348255/)]

## Abbreviations

**mHealth:** mobile health

**OR:** odds ratio

**TMI:** text messaging-based interaction

*Edited by L Buis; submitted 02.09.20; peer-reviewed by M Kang, S Badawy, S Patil; comments to author 30.10.20; revised version received 04.01.21; accepted 23.02.21; published 01.04.21.*

*Please cite as:*

Linnemayr S, Zutshi R, Shadel W, Pedersen E, DeYoreo M, Tucker J

*Text Messaging Intervention for Young Smokers Experiencing Homelessness: Lessons Learned From a Randomized Controlled Trial*  
*JMIR Mhealth Uhealth* 2021;9(4):e23989

URL: <https://mhealth.jmir.org/2021/4/e23989>

doi: [10.2196/23989](https://doi.org/10.2196/23989)

PMID: [33792551](https://pubmed.ncbi.nlm.nih.gov/33792551/)

©Sebastian Linnemayr, Rushil Zutshi, William Shadel, Eric Pedersen, Maria DeYoreo, Joan Tucker. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 01.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Gender Differences in Satisfaction With a Text Messaging Program (Text4Hope) and Anticipated Receptivity to Technology-Based Health Support During the COVID-19 Pandemic: Cross-sectional Survey Study

Reham Shalaby<sup>1</sup>, MD; Wesley Vuong<sup>2</sup>, MPH; Marianne Hrabok<sup>3</sup>, PhD; April Gusnowski<sup>2</sup>, BA; Kelly Mrklas<sup>4,5</sup>, MSc; Daniel Li<sup>2</sup>, MSc, MD; Mark Snaterse<sup>2</sup>, BSc; Shireen Surood<sup>2</sup>, PhD; Bo Cao<sup>1</sup>, PhD; Xin-Min Li<sup>1</sup>, MD, PhD; Russell Greiner<sup>1</sup>, PhD; Andrew James Greenshaw<sup>1</sup>, PhD; Vincent Israel Opoku Agyapong<sup>1</sup>, MD, PhD

<sup>1</sup>Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

<sup>2</sup>Addiction and Mental Health, Alberta Health Services, Edmonton, AB, Canada

<sup>3</sup>Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

<sup>4</sup>Strategic Clinical Networks, Provincial Clinical Excellence, Alberta Health Services, Calgary, AB, Canada

<sup>5</sup>Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Edmonton, AB, Canada

**Corresponding Author:**

Vincent Israel Opoku Agyapong, MD, PhD

Department of Psychiatry

Faculty of Medicine and Dentistry

University of Alberta

8440 112 St NW

Edmonton, AB, T6G 2B7

Canada

Phone: 1 7807144315

Email: [agyapong@ualberta.ca](mailto:agyapong@ualberta.ca)

## Abstract

**Background:** In March 2020, Text4Hope—a community health service—was provided to Alberta residents. This free service aims to promote psychological resilience and alleviate pandemic-associated stress, anxiety, and depression symptoms during the COVID-19 pandemic.

**Objective:** This study aimed to evaluate the feedback, satisfaction, experience, and perceptions of Text4Hope subscribers and to examine any differences based on gender after subscribers received 6 weeks of daily supportive text messages. Additionally, this study examined subscribers' anticipated receptivity to technology-based medical services that could be offered during major crises, emergencies, or pandemics.

**Methods:** Individuals self-subscribed to Text4Hope to receive daily supportive text messages for 3 months. Subscribers were invited to complete a web-based survey at 6 weeks postintervention to provide service satisfaction-related information. Overall satisfaction was assessed on a scale of 0-10, and satisfaction scores were analyzed using a related-measures *t* test. Likert scale satisfaction responses were used to assess various aspects of the Text4Hope program. Gender differences were analyzed using one-way analysis of variance (ANOVA) and Chi-square analyses.

**Results:** A total of 2032 subscribers completed the baseline and 6-week surveys; 1788 (88%) were female, 219 (10.8%) were male, and 25 (1.2%) were other gender. The mean age of study participants was 44.58 years (SD 13.45 years). The mean overall satisfaction score was 8.55 (SD 1.78), suggesting high overall satisfaction with Text4Hope. The ANOVA analysis, which was conducted using the Welch test ( $n=1716$ ), demonstrated that females had significantly higher mean satisfaction scores than males (8.65 vs 8.11, respectively; mean difference=0.546; 95% CI 0.19 to 0.91;  $P<.001$ ) and nonsignificantly lower satisfaction scores than other gender respondents (mean difference=-0.938; 95% CI -0.37 to 2.25;  $P=.15$ ). More than 70% of subscribers agreed that Text4Hope helped them cope with stress (1334/1731, 77.1%) and anxiety (1309/1728, 75.8%), feel connected to a support system (1400/1729, 81%), manage COVID-19-related issues (1279/1728, 74%), and improve mental well-being (1308/1731, 75.6%). Similarly, subscribers agreed that messages were positive, affirmative, and succinct. Messages were always or often read by 97.9% (1681/1716) of respondents, and more than 20% (401/1716, 23.4%) always or often returned to messages. The majority

of subscribers (1471/1666, 88.3%) read the messages and either reflected upon them or took a positive action. Subscribers welcomed almost all technology-based services as part of their health care during crisis or emergency situations. Text4Hope was perceived to be effective by many female subscribers, who reported higher satisfaction and improved coping after receiving text messages for 6 weeks.

**Conclusions:** Respondents affirmed the high quality of the text messages with their positive feedback. Technology-based services can provide remotely accessible and population-level interventions that align with the recommended physical distancing practices for pandemics. Text4Hope subscriber feedback revealed high satisfaction and acceptance at 6 weeks postintervention.

**International Registered Report Identifier (IRRID):** RR2-10.2196/19292

(*JMIR Mhealth Uhealth* 2021;9(4):e24184) doi:[10.2196/24184](https://doi.org/10.2196/24184)

## KEYWORDS

COVID-19; Text4Hope; satisfaction; mobile phone; text; anxiety; depression; stress, pandemic; e-mental health; gender

## Introduction

### Background

On March 11, 2020, the World Health Organization declared COVID-19 a global pandemic [1]. By March 23, 2020, there were 332,930 COVID-19 cases worldwide and 14,509 deaths attributed to the pandemic [2]. On this date, Alberta Health Services (the provincial health authority in Alberta, Canada) launched Text4Hope—a free, mobile, community mental health service that aims to support mental well-being and resilience, improve coping mechanisms, and safeguard against pandemic-associated thoughts in Alberta residents [3]. The service was advertised on the Alberta Health Services and Text4Hope funders' websites and was launched on March 23, 2020. Thousands of people have signed up for the service, and enrollment continues to increase to date. Text4Hope is a text-based mental health support program that involves daily, evidence-based, cognitive behavioral therapy–derived text messages. These messages were carefully designed to accompany a rapidly evolving health crisis and to be scalable, remotely deliverable, and accessible. They were also designed to be cost-effective for funding organizations and free to subscribers [4]. The Text4Hope program was developed based on lessons from the Text4Mood and Text4Support programs [5,6]. Similar to the Text4Mood program, individual Text4Hope self-subscribers receive daily text messages. However, while the Text4Mood messages were crafted to mainly address anxiety, depression, and general well-being among residents of Northern Alberta, the Text4Hope messages were crafted to predominantly address COVID-19–related stress, anxiety, and depression among all Albertans. In contrast to both Text4Mood and Text4Hope, Text4Support was specifically designed to provide support for the eight most commonly observed addiction and mental health concerns in the Edmonton Zone [6]. In this program, a mental health therapist or psychiatrist sorts clients into 1 of the 8 categories, and patients are enrolled by a coordinator inputting the patients' mobile phone numbers into a web-based program. Text4Hope fills a service gap in Alberta, as social distancing measures may have resulted in high-risk individuals (from a health perspective) not being able to access addiction and mental health services during the early stage of the pandemic. Text4Hope also offers mental health support to those who might not feel comfortable with in-person contact.

During similar crises, the effective and efficient mobilization of community resources was strongly encouraged to support and properly meet mental health needs and avoid future adverse mental health consequences [7]. During pandemics, negative thoughts accompanied by growing uncertainties can pose a threat to personal health and mental well-being. The transmissibility of SARS-CoV-2 has been shown to exceed that of similar viruses (eg, MERS-CoV [Middle East respiratory syndrome coronavirus], H1N1, and SARS-CoV [severe acute respiratory syndrome coronavirus]) [8]. As such, strict policies and regulations were enforced to contain viral spread, including physical distancing, self-isolation, quarantine, travel restrictions, the closure of public schools, and disinfection protocols. However, these measures have likely contributed to mental strain and psychological distress during the COVID-19 pandemic [9,10]. Other iterations of texting programs were developed to support patients with major depressive disorders [11] and alcohol use disorder [12,13]. Individuals in these programs reported an improvement in depression scores and felt better supported in their attempts to quit drinking alcohol after receiving text messages [12,14]. Supportive text messaging services can be tailored to meet the needs of diverse populations. For example, Text4baby and Quit4baby are two services that are provided to pregnant women in the United States [15,16], while Text4Mood and Text4Support are mental health services that are provided to people in Canada [5,6]. Ultimately, such services provide people with hopefulness and support and aim to close the psychological treatment gap in health care systems [5].

To make the best use of resources and enhance the use of texting technology as part of routine practice in health care, it is essential to assess user satisfaction and better understand subscribers' experiences. The assessment of user satisfaction is a quality method that affects client retention and clinical outcomes [17]. In the customer service industry, relative satisfaction and customer expectations are considered critical components for guaranteeing customer loyalty [18]. In health care systems, self-reported continuity of care strongly correlates with client satisfaction. A recent study has demonstrated that a 7.2% reduction in the frequency of reporting “at least good overall satisfaction” was associated with a 1% increase in hospital bed occupancy [17]. Generally, asynchronous web-based and text-based services have been accepted by an increasing number of individuals who perceive such services

as supportive and promising [19]. Most of these programs have usually stated that more than 85% of text message recipients report high satisfaction, high convenience, easy use, and better control over life activities, while above 90% report increased life productivity after receiving text messages [20,21]. Additionally, telephone services are frequently associated with having lower attrition rates than face-to-face services, which is likely due to the accessibility provided by technology that removes geographical barriers. This is especially helpful to those who are tentative about seeking medical attention or require medications [22]. Agyapong and colleagues [5], who evaluated Text4Mood, found that 80% of participants agreed that asynchronous supportive text messages should be provided during follow-up care, and approximately 50% of participants agreed to the use of videoconferencing consultations. A number of variables may affect users' satisfaction with texting services, such as sociodemographic characteristics, health status, and disease severity. Similarly, one's gender identity may be an important determinant of service acceptability and satisfaction. However, it should be noted that inconsistent findings have been reported for gender identity effects. Although females are highly accepting of surveys and have a high desire to respond to surveys that are delivered to them via a texting service [23], in a feasibility study, the high fidelity of a texting service program was also reported when the program was provided to a group of disadvantaged men at risk of substance or alcohol abuse [24]. In yet another study, authors found no difference between male and female university students in terms of their satisfaction with texting services for alcohol use intervention [14]. Additionally, the initial reports of our program revealed that a majority of our subscribers reported their gender as female (86.9%). This overrepresentation of females in text messaging services has necessitated investigations into user satisfaction and anticipated agreement to receiving technology-based medical services based on gender. Such investigations will allow targeted gender-based interventions to be developed in accordance with user preferences.

This study occurred in Alberta, the Canadian Province where the Text4Hope program was launched. As of July 1, 2020, Alberta had a population of 4,421,876 people, with 68% of the population aged between 15 and 64 years. Alberta has consistently consisted of more males than females (101 males per 100 females), mainly due to the large proportion of working-age males migrating to the province [25]. In 2006, the racial and ethnic composition of Alberta was 80.3% White Canadians, 13.9% visible minority groups, and 5.8% Indigenous groups (3% First Nations, 2.6% Metis, and 0.1% other Indigenous groups). Visible minority groups included the following: Chinese (3.7%), South Asian (3.2%), Filipino (1.6%), Black (1.4%), Southeast Asian (0.9%), Latin American (0.8%), Arab (0.8%), Korean (0.4%), West Asian (0.3%), and Japanese 0.3% [26]. In 2016, more than half (54%) of Canadians aged 25-64 years had either college or university qualifications (an increase from the 48.3% in 2006) [27]. Alberta's gross domestic product at basic prices was CAN \$334.5 billion (US \$265.2 billion) in 2019 (largely unchanged from Alberta's gross domestic product in 2018) [28].

## Objective

The aim of this study was to evaluate subscribers' overall satisfaction with Text4Hope; obtain feedback about subscribers' experiences and the impact of the texting intervention; explore the perceptions of subscribers about their anticipated receptivity toward diverse, technology-based medical services that are offered as a part of their health care during major crises, emergencies, or pandemics (such as the COVID-19 pandemic); and examine any differences that are based on gender after subscribers received 6 weeks of daily supportive text messages.

## Hypotheses

Based on previous Text4Mood research [5], our hypotheses were as follows: (1) the mean overall satisfaction level with Text4Hope would be at least 7.5 (75%) and (2) at least 75% of subscribers would express anticipated agreement with receiving diverse, technology-based medical services during crises or emergencies. Additionally, we believed that there would be a difference in the satisfaction measure based on the self-declared gender identity of the respondents.

## Methods

### Study Design

This cross-sectional study assessed subscribers' satisfaction and experiences with Text4Hope and their perceptions of technology-based support after they received 6 weeks of daily text messages.

### Data Collection

The data collection methods were fully described in the study protocol [29]. In summary, subscribers joined the Text4Hope program [3] and received daily supportive text messages for 3 months by texting the word "COVID19HOPE" to a short code number. The messages were in line with a cognitive behavioral framework that addressed the aspects of potential stresses, anxiety, and depression, and the content was written by mental health professionals. Text message delivery was unidirectional and not specifically tailored to the end users. The following are examples of the messages that were sent:

*When bad things happen that we can't control, we often focus on the things we can't change. Focus on what you can control; what you can do to help yourself (or someone else) today. [Example 1]*

*What lies behind you and what lies before you are tiny matters compared to what lies within you. Have faith in yourself and success can be yours. [Example 2]*

*Set goals for today, even if they are small. Goals should be "SMART": Specific, Measurable, Achievable, Realistic, and Timely. [Example 3]*

The messages were uploaded to a web-based platform, which delivered automated messages at 9 AM. The first message welcomed subscribers to the service and invited them to voluntarily complete a web-based baseline survey, which was used to capture demographic and clinical information that primarily pertained to anxiety, stress, depression, and

self-isolation. At 6 weeks postintervention, subscribers were invited (via a text message link) to complete a follow-up web-based survey.

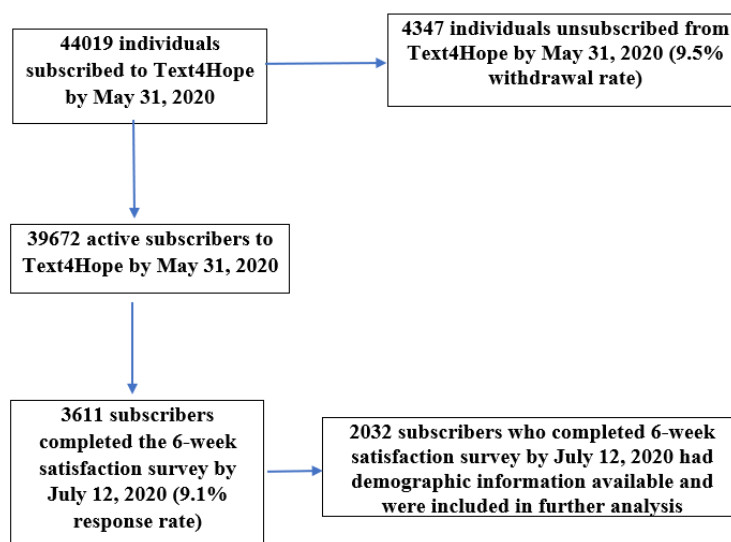
The 6-week survey included standardized scales that were used for the Text4Hope baseline assessments [30,31] as well as an adopted version of the Text4Mood user satisfaction survey [5]. Each survey took 5-10 minutes to complete. No incentives were offered to respondents for completing the baseline or 6-week surveys. Consent was implied if participants clicked on the survey links and submitted their responses.

Participation in the program was voluntary, and the receipt of supportive text messages was not contingent on survey completion. Subscribers could opt out of Text4Hope at any time by texting the word "STOP" to a short code number.

Six-week satisfaction data were collected between May 31 and July 12, 2020. Figure 1 depicts a subscriber flowchart, which indicates the number of subscribers who completed the web-based surveys at each time point.

The study protocol [29] was approved by the Research and Ethics Board of the University of Alberta (approval number: Pro00086163).

**Figure 1.** Subscription flowchart.



## Outcome Measures

The primary outcome measure was subscribers' overall satisfaction with the Text4Hope daily supportive text messages. Overall satisfaction at 6 weeks postintervention was based on an 11-point Likert scale (0=very dissatisfied; 5=neither satisfied nor dissatisfied; 10=very satisfied). This overall satisfaction score allowed us to determine whether people liked texting-based services. If people are satisfied with the population-based services they receive, then the services are potentially feasible and can aid in future service planning during pandemics. The satisfaction scale has been used to compare service satisfaction across all addiction and mental health services in the Edmonton Zone. The reliability and validity of this scale has not been tested, although it has been in use for several years.

Secondary outcomes included the perceived impacts of and subscribers' feedback for the daily supportive text messages at 6 weeks postintervention as well as subscribers' anticipated receptivity to diverse, technology-based medical services (eg, telephone, videoconferencing, and email for health care) during the COVID-19 pandemic. Gender differences in both primary and secondary measures constituted the exploratory outcome measures.

## Sample Size Considerations

In total, 44,019 individuals were subscribed to Text4Hope in May 31, 2020. We estimated that a sample size of 1775 was needed to estimate the overall mean satisfaction rate (based on an 11-point scale from 0 to 10) for the entire population with a 3% margin of error and 99% confidence.

## Analysis

Data were analyzed using SPSS Statistics for Windows, version 26 (IBM Corporation) [32]. Demographic characteristics were summarized as raw numbers and percentages. We measured subscribers' overall satisfaction on an 11-point Likert scale (0=very dissatisfied; 5=neither satisfied nor dissatisfied; 10=very satisfied) and analyzed responses by using the related sample *t* test. We explored gender differences in satisfaction, which was measured on the same scale, by using one-way analysis of variance (ANOVA) tests. A Bonferroni-corrected, two-tailed criterion ( $\alpha < .002$ ) was used to determine statistical differences. Likert scale satisfaction responses to various aspects of Text4Hope and anticipated receptivity to technology-based interventions (web-based counseling, telephone counseling, text and email messaging, telephone consultations for physical and mental health, and video consultations for physical and mental health) were summarized as frequency counts of response categories and percentages. We compared gender differences in satisfaction and preferences for technology-based interventions by using the Fisher exact test with two-tailed,

Bonferroni-corrected criteria for 23 variables ( $\alpha < .002$ ) to determine statistical differences. There was no imputation for missing data, and the results were based on completed survey responses.

Between May 31 and July 12, 2020, 39,672 active Text4Hope subscribers were invited to complete the 6-week survey. Of

these subscribers, 3611 completed the survey, yielding a response rate of 9.1%. Of the 2032 subscribers who had available demographic information from their baseline survey and were included in further analysis, 1788 (88%) were female, 219 (10.8%) were male, and 25 (1.2%) were other gender. [Table 1](#) provides a descriptive analysis of the demographics of respondents.

**Table 1.** Demographic and clinical characteristics of respondents at 6 weeks postintervention.

Variables	Male (n=219), n (%)	Female (n=1788), n (%)	Other gender (n=25), n (%)	Overall (N=2032), n (%)
<b>Age (years)</b>				
≤25	15 (7)	151 (8.6)	7 (28)	173 (8.6)
26-40	53 (24.8)	490 (27.7)	12 (48)	555 (27.7)
41-60	105 (49.1)	891 (50.5)	4 (16)	1000 (49.9)
>60	41 (19.2)	234 (13.3)	2 (8)	277 (13.8)
<b>Ethnicity</b>				
White	177 (81.2)	1492 (83.9)	19 (76)	1688 (83.5)
Indigenous	6 (2.8)	54 (3)	0 (0)	60 (3)
Asian	15 (6.9)	90 (5.1)	1 (4)	106 (5.2)
Other	20 (9.2)	142 (8)	5 (20)	167 (8.3)
<b>Education</b>				
Less than a high school diploma	7 (4)	35 (2.3)	2 (8.7)	44 (2.6)
High school diploma	14 (8)	102 (6.8)	1 (4.3)	117 (6.9)
Postsecondary education	155 (88.1)	1349 (90.2)	19 (82.6)	1523 (89.9)
Other education	0 (0)	9 (0.6)	1 (4.3)	10 (0.6)
<b>Employment status</b>				
Employed	120 (69)	1059 (71.5)	12 (52.2)	1191 (70.9)
Unemployed	26 (14.9)	177 (11.9)	3 (13)	206 (12.3)
Retired	23 (13.2)	151 (10.2)	2 (8.7)	176 (10.5)
Student	4 (2.3)	71 (4.8)	5 (21.7)	80 (4.8)
Other	1 (0.6)	24 (1.6)	1 (4.3)	26 (1.5)
<b>Relationship status</b>				
Married/cohabiting/partnered	112 (63.3)	987 (66)	11 (47.8)	1110 (65.5)
Separated/divorced	14 (7.9)	154 (10.3)	1 (4.3)	169 (10)
Widowed	3 (1.7)	37 (2.5)	1 (4.3)	41 (2.4)
Single	46 (26)	303 (20.3)	9 (39.1)	358 (21.1)
Other	2 (1.1)	14 (0.9)	1 (4.3)	17 (1)
<b>Housing status</b>				
Own home	122 (69.7)	1037 (70.1)	12 (52.2)	1171 (69.8)
Living with family	14 (8)	132 (8.9)	4 (17.4)	150 (8.9)
Renting	38 (21.7)	300 (20.3)	5 (21.7)	343 (20.5)
Other	1 (0.6)	10 (0.7)	2 (8.7)	13 (0.8)

## Results

### Demographic and Clinical Characteristics

Table 1 displays subscribers' demographic characteristics based on different genders. The data indicated that most respondents were aged between 26 and 60 years (1555/2032, 77.6%); were White (1688/2032, 83.5%); were married, cohabiting, or partnered (1110/2032, 65.5%); reported the completion of postsecondary education (1523/2032, 89.9%); were employed (1191/2032, 70.9%); and owned their own home (1171/2032, 69.8%).

### Primary Outcome Measure

Respondents were asked to rate their overall satisfaction with the daily supportive text messaging (Text4Hope) service on a scale of 0-10, in which 0 represented "very dissatisfied," 5

represented "neither satisfied nor dissatisfied," and 10 represented "very satisfied." Respondents' (n=2940) mean overall satisfaction score was 8.55 (SD 1.78), suggesting that overall, respondents' satisfaction with the Text4Hope program was high. The ANOVA analysis, which was conducted using the Welch test (n=1716), demonstrated that females had significantly higher mean satisfaction scores than males (8.65 vs 8.11, respectively; mean difference=0.546; 95% CI 0.19 to 0.91;  $P<.001$ ) and nonsignificantly lower satisfaction scores than other gender respondents (mean difference=-0.938; 95% CI -0.37 to 2.25;  $P=0.15$ ).

### Secondary Outcome Measures

In Table 2, we show subscribers' level of agreement regarding Text4Hope benefits. This table displays the perceived impact of Text4Hope messages after subscribers received daily text messages for 6 weeks.

**Table 2.** Gender differences in the perceived impact of daily messages at 6 weeks postintervention.

Perceived impact of daily messages from Text4Hope	Male, n (%)	Female, n (%)	Other gender, n (%)	<i>P</i> value <sup>a</sup>	Total, n (%)
<b>Helped subscribers cope with stress related to the COVID-19 pandemic</b>					
Agree	144 (75.8)	1177 (77.4)	13 (61.9)	.05	1334 (77.1)
Neutral	33 (17.4)	284 (18.7)	5 (23.8)	N/A <sup>b</sup>	322 (18.6)
Disagree	13 (6.8)	59 (3.9)	3 (14.3)	N/A	75 (4.3)
<b>Helped subscribers cope with anxiety related to the COVID-19 pandemic</b>					
Agree	133 (70.4)	1162 (76.5)	14 (66.7)	.05	1309 (75.8)
Neutral	44 (23.3)	297 (19.6)	4 (19)	N/A	345 (20)
Disagree	12 (6.3)	59 (3.9)	3 (14.3)	N/A	74 (4.3)
<b>Helped subscribers cope with depression related to the COVID-19 pandemic</b>					
Agree	103 (54.5)	856 (56.4)	9 (42.9)	.04	968 (56.1)
Neutral	63 (33.3)	561 (37)	9 (42.9)	N/A	633 (36.7)
Disagree	23 (12.2)	100 (6.6)	3 (14.3)	N/A	126 (7.3)
<b>Helped subscribers cope with loneliness related to the COVID-19 pandemic</b>					
Agree	71 (37.4)	757 (49.9)	9 (42.9)	.01	837 (48.5)
Neutral	85 (44.7)	592 (39.1)	9 (42.9)	N/A	686 (39.7)
Disagree	34 (17.9)	167 (11)	3 (14.3)	N/A	204 (11.8)
<b>Made subscribers feel connected to a support system during the COVID-19 pandemic</b>					
Agree	144 (75.8)	1242 (81.8)	14 (66.7)	.05	1400 (81)
Neutral	33 (17.4)	211 (13.9)	4 (19)	N/A	248 (14.3)
Disagree	13 (6.8)	65 (4.3)	3 (14.3)	N/A	81 (4.7)
<b>Made subscribers feel hopeful about managing issues related to the COVID-19 pandemic</b>					
Agree	134 (70.5)	1133 (74.7)	12 (57.1)	.09	1279 (74)
Neutral	46 (24.2)	324 (21.4)	6 (28.6)	N/A	376 (21.8)
Disagree	10 (5.3)	60 (4)	3 (14.3)	N/A	73 (4.2)
<b>Improved subscribers' overall mental well-being</b>					
Agree	136 (71.6)	1159 (76.2)	13 (61.9)	.06	1308 (75.6)
Neutral	38 (20)	289 (19)	5 (23.8)	N/A	332 (19.2)
Disagree	16 (8.4)	72 (4.7)	3 (14.3)	N/A	91 (5.3)
<b>Enhanced subscribers' quality of life</b>					
Agree	104 (55)	941 (62.5)	12 (60)	.11	1057 (61.7)
Neutral	68 (36)	474 (31.5)	5 (25)	N/A	547 (31.9)
Disagree	17 (9)	90 (6)	3 (15)	N/A	110 (6.4)

<sup>a</sup>Bonferroni-corrected, two-tailed criteria for significance ( $\alpha < .002$ ).

<sup>b</sup>N/A: not applicable.

The results in Table 2 indicate that about three-quarters of respondents agreed that the daily text messages helped them cope with stress (1334/1731, 77.1%) and anxiety (1309/1728, 75.8%) as well as manage COVID-19-related issues (1279/1728, 74%), while about half of the respondents agreed that the messages helped them cope with depression (968/1727, 56.1%) and loneliness (837/1727, 48.5%). About 80% of respondents agreed that they felt connected to a support system due to receiving the daily messages (1400/1729, 81%), a little over

70% of respondents agreed that the daily messages helped to improve their mental well-being (1308/1731, 75.6%), and about 60% of respondents agreed that the daily messages helped to enhance their quality of life (1057/1714, 61.7%). Overall, compared to males and respondents of other gender identities, a higher proportion of females agreed with all Text4Hope benefits; however, there were no statistically significant gender differences in the levels of agreement expressed for all areas assessed.



**Table 3** describes subscribers' opinions about Text4Hope messages after they received 6 weeks of daily text messages. The data indicated that about three-quarters of respondents always found the Text4Hope text messages to be positive (1336/1732, 77.1%), affirmative (1231/1727, 71.3%), and succinct (1254/1722, 72.8%). More than 80% of respondents (1505/1753, 87.4%) indicated that the messages were always or often relevant. Again, compared to males and respondents of other gender identities, a higher proportion of females reported that they found the messages to be always positive, affirmative, succinct, and relevant ( $P < .001$  for each posthoc comparison using z-scores).

Most respondents (1531/1716, 89.2%) indicated that they always read the text messages, and about 20% of respondents indicated

that they always or often returned to read the text messages (401/1716, 23.4%). Neither factor indicated gender differences upon analysis. **Table 3** data shows that slightly more than 70% respondents (1270/1666, 76.2%) indicated that they read and reflected on the text messages, while about 10% of respondents indicated that they took positive or beneficial actions after reading the text messages (201/1666, 12.1%). Although not statistically significant ( $P = .003$ ), compared to males and respondents of other gender identities, a higher proportion of females indicated that they read the text messages, reflected on the messages, and took positive or beneficial actions after reading the messages. No subscribers indicated that they read the messages and took a negative action.

**Table 3.** Gender differences in the feedback about Text4Hope messages at 6 weeks postintervention.

Feedback	Male, n (%)	Female, n (%)	Other gender, n (%)	P value <sup>a</sup>	Total, n (%)
<b>Text4Hope text messages were positive</b>					
Always	131 (68.9)	1193 (78.4)	12 (57.1)	<.001	1336 (77.1)
Often	55 (28.9)	291 (19.1)	6 (28.6)	N/A <sup>b</sup>	352 (20.3)
Sometimes	3 (1.6)	35 (2.3)	2 (9.5)	N/A	40 (2.3)
Rarely	1 (0.5)	2 (0.1)	1 (4.8)	N/A	4 (0.2)
Never	0 (0)	0 (0)	0 (0)	N/A	0 (0)
<b>Text4Hope text messages were affirmative</b>					
Always	118 (62.8)	1104 (72.7)	9 (42.9)	<.001	1231 (71.3)
Often	57 (30.3)	347 (22.9)	10 (47.6)	N/A	414 (24)
Sometimes	11 (5.9)	58 (3.8)	1 (4.8)	N/A	70 (4.1)
Rarely	1 (0.5)	8 (0.5)	0 (0)	N/A	9 (0.5)
Never	1 (0.5)	1 (0.1)	1 (4.8)	N/A	3 (0.2)
<b>Text4Hope text messages were succinct</b>					
Always	128 (67.7)	1114 (73.7)	12 (57.1)	.09	1254 (72.8)
Often	49 (25.9)	300 (19.8)	5 (23.8)	N/A	354 (20.6)
Sometimes	12 (6.3)	92 (6.1)	4 (19)	N/A	108 (6.3)
Rarely	0 (0)	6 (0.4)	0 (0)	N/A	6 (0.3)
Never	0 (0)	0 (0)	0 (0)	N/A	0 (0)
<b>Text4Hope text messages were relevant</b>					
Always	95 (50.5)	945 (62.4)	12 (57.1)	<.001	1052 (61.1)
Often	64 (34)	386 (25.5)	3 (14.3)	N/A	453 (26.3)
Sometimes	20 (10.6)	163 (10.8)	3 (14.3)	N/A	186 (10.8)
Rarely	7 (3.7)	19 (1.3)	2 (9.5)	N/A	28 (1.6)
Never	2 (1.1)	1 (0.1)	1 (4.8)	N/A	4 (0.2)
<b>Subscribers' frequency of reading messages</b>					
Always	161 (84.7)	1351 (89.8)	19 (90.5)	.61	1531 (89.2)
Often	23 (12.1)	125 (8.3)	2 (9.5)	N/A	150 (8.7)
Sometimes	5 (2.6)	25 (1.7)	0 (0)	N/A	30 (1.7)
Rarely	1 (0.5)	2 (0.1)	0 (0)	N/A	3 (0.2)
Never	0 (0)	2 (0.1)	0 (0)	N/A	2 (0.1)
<b>Subscribers' frequency of returning to messages</b>					
Always	7 (3.7)	73 (4.9)	0 (0)	.47	80 (4.7)
Often	33 (17.4)	287 (19.1)	1 (4.8)	N/A	321 (18.7)
Sometimes	76 (40)	635 (42.2)	13 (61.9)	N/A	724 (42.2)
Rarely	46 (24.2)	327 (21.7)	4 (19)	N/A	377 (22)
Never	28 (14.7)	183 (12.2)	3 (14.3)	N/A	214 (12.5)
<b>Actions taken by subscribers after reading text messages</b>					
Read text and took a positive or beneficial action	14 (7.7)	186 (12.7)	1 (5)	.003	201 (12.1)
Read text and reflected on the messages	138 (75.4)	1119 (76.5)	13 (65)	N/A	1270 (76.2)
Read the text and took no action	25 (13.7)	138 (9.4)	6 (30)	N/A	169 (10.1)

Feedback	Male, n (%)	Female, n (%)	Other gender, n (%)	<i>P</i> value <sup>a</sup>	Total, n (%)
Read text and took a negative or harmful action	0 (0)	0 (0)	0 (0)	N/A	0 (0)
Did not read the text	2 (1.1)	2 (0.1)	0 (0)	N/A	4 (0.2)
Other	4 (2.2)	18 (1.2)	0 (0)	N/A	22 (1.3)

<sup>a</sup>Bonferroni-corrected, two-tailed criteria for significance ( $\alpha < .002$ ).

<sup>b</sup>N/A: not applicable.

We explored subscribers' anticipated receptivity to welcoming diverse, technology-based services as part of their health care during crisis or emergency situations, such as the COVID-19 pandemic. The results displayed in Table 4 suggest that at least 80% of respondents agreed with receiving web-based counseling (1390/1674, 83%), telephone counseling (1346/1672, 80.5%), and text messages (1465/1669, 87.8%) as part of their health care during any crisis or emergency situation, such as the COVID-19 pandemic. There were no gender differences in respondents' preferences for welcoming web-based counseling, telephone counseling, and text messaging as part of their health care during any crisis or emergency situation. Similarly, about 70% of respondents agreed with receiving consultations via

video and telephone for both physical (video: 1190/1674, 71.1%; telephone: 1193/1665, 71.7%) and mental (video: 1244/1674, 74.3%; telephone: 1245/1669, 74.6%) health care during any crisis or emergency situation, such as the COVID-19 pandemic. There were no gender-based differences in expressed preferences. Finally, about 60% of respondents agreed with receiving email messages as part of their health care during a crisis or emergency situation, such as the COVID-19 pandemic (1084/1669, 64.9%). Compared to female and male respondents, a higher proportion of other gender respondents agreed with receiving email messages as part of their health care during a crisis or emergency situation.

**Table 4.** Anticipated receptivity of subscribers to receiving diverse, technology-based services as part of their health care during crisis or emergency situations, such as the COVID-19 pandemic.

Subscribers' anticipated receptivity to services	Male, n (%)	Female, n (%)	Other gender, n (%)	P value <sup>a</sup>	Total, n (%)
<b>Subscribers would welcome web-based counseling for stress, anxiety, and depression</b>					
Agree	152 (80.9)	1220 (83.3)	18 (85.7)	.55	1390 (83)
Neutral	26 (13.8)	198 (13.5)	3 (14.3)	N/A <sup>b</sup>	227 (13.6)
Disagree	10 (5.3)	47 (3.2)	0 (0)	N/A	57 (3.4)
<b>Subscribers would welcome telephone counseling for stress, anxiety, and depression</b>					
Agree	151 (80.3)	1176 (80.4)	19 (90.5)	.80	1346 (80.5)
Neutral	29 (15.4)	229 (15.7)	2 (9.5)	N/A	260 (15.6)
Disagree	8 (4.3)	58 (4)	0 (0)	N/A	66 (3.9)
<b>Subscribers would welcome text messaging for stress, anxiety, and depression</b>					
Agree	159 (84.6)	1288 (88.2)	18 (85.7)	.12	1465 (87.8)
Neutral	19 (10.1)	132 (9)	1 (4.8)	N/A	152 (9.1)
Disagree	10 (5.3)	40 (2.7)	2 (9.5)	N/A	52 (3.1)
<b>Subscribers would welcome email messaging for stress, anxiety, and depression</b>					
Agree	106 (56.7)	962 (65.8)	16 (76.2)	.01	1084 (64.9)
Neutral	45 (24.1)	345 (23.6)	4 (19)	N/A	394 (23.6)
Disagree	36 (19.3)	154 (10.5)	1 (4.8)	N/A	191 (11.4)
<b>Subscribers would welcome mental health video consultations</b>					
Agree	132 (70.2)	1094 (74.7)	18 (85.7)	.29	1244 (74.3)
Neutral	42 (22.3)	284 (19.4)	1 (4.8)	N/A	327 (19.5)
Disagree	14 (7.4)	87 (5.9)	2 (9.5)	N/A	103 (6.2)
<b>Subscribers would welcome physical health video consultations</b>					
Agree	119 (63.3)	1055 (72)	16 (76.2)	.12	1190 (71.1)
Neutral	50 (26.6)	279 (19)	4 (19)	N/A	333 (19.9)
Disagree	19 (10.1)	131 (8.9)	1 (4.8)	N/A	151 (9)
<b>Subscribers would welcome mental health telephone consultations</b>					
Agree	126 (67.4)	1102 (75.4)	17 (81)	.19	1245 (74.6)
Neutral	44 (23.5)	259 (17.7)	3 (14.3)	N/A	306 (18.3)
Disagree	17 (9.1)	100 (6.8)	1 (4.8)	N/A	118 (7.1)
<b>Subscribers would welcome physical health telephone consultations</b>					
Agree	124 (66.3)	1052 (72.2)	17 (81)	.30	1193 (71.7)
Neutral	44 (23.5)	258 (17.7)	3 (14.3)	N/A	305 (18.3)
Disagree	19 (10.2)	147 (10.1)	1 (4.8)	N/A	167 (10)

<sup>a</sup>Bonferroni-corrected, two-tailed criteria for significance ( $\alpha < .002$ ).

<sup>b</sup>N/A: not applicable.

## Discussion

This study provided results regarding subscribers' satisfaction with Text4Hope after they received the texting intervention for 6 weeks. Our results revealed considerable satisfaction with Text4Hope. The total number of subscribers who completed the baseline and 6-week surveys was 2032, and a majority of subscribers were female (1788/2032, 88%). The mean age of study participants was 44.58 years. Overall service satisfaction

was high, and more than 70% of subscribers agreed that Text4Hope helped them cope with stress (1334/1731, 77.1%) and anxiety (1309/1728, 75.8%), feel connected to a support system (1400/1729, 81%), manage COVID-19-related issues (1279/1728, 74%), and improve mental well-being (1308/1731, 75.6%). Similarly, subscribers agreed that the text messages were positive, affirmative, and succinct. Text messages were always or often read by 97.9% (1681/1716) of respondents, and more than 20% (401/1716, 23.4%) always or often returned to

messages. Most subscribers (1471/1666, 88.3%) read the messages and either reflected upon them or took a positive action. Subscribers welcomed almost all technology-based services as part of their health care during crisis or emergency situations. Text4Hope was perceived to be effective by more female subscribers than male or other gender subscribers. The withdrawal rate for Text4Hope was approximately 10% at 6 weeks postintervention. Untailored and unilateral texting services often have high withdrawal rates that range from 0% to 57% [14,33]. Additionally, prior studies have reported that withdrawal rates may be higher for people who receive interventions via SMS text messages compared to those for people who receive the same intervention via email [14]. In a review of 93 mental health apps that target anxiety, depression, or emotional well-being, the median 15-day and 30-day app retention rates were only 3.9% (IQR 10.3%) and 3.3% (IQR 6.2%), respectively [34]. It is possible that our Text4Hope program achieved a higher retention rate compared to those of other mental health apps because it is unidirectional and requires no additional effort or action on the part of the subscriber following enrollment. It is also possible that the message content, which was crafted by mental health professionals; the high anxiety, stress, and depression levels that the population has experienced due to the COVID-19 pandemic; and the reduced availability of face-to-face services contributed to the high Text4Hope retention rate.

Female respondents comprised the majority of the sample in our study (1788/2032, 88%). In other texting-based services, females were also highly represented (>80% of participants) [5]. There were obvious gender differences in subscriber satisfaction rates for Text4Hope. Another study, in which 240 university students received a fully automated, multiple-session alcohol intervention, reported that the majority of students were satisfied with the content and length of the texts; no gender-based differences in responses were reported [14].

Subscribers' overall satisfaction with our provided service (8.55) was high. This is in line with the 95% satisfaction rate of the Text4Mood program reported by Agyapong et al [5]. Similar findings were reported in a review of text message use among a population with mental health concerns [35]. Bendsten and Bendsten [14] previously reported on participant satisfaction (range 57.9%-84.6%) in relation to the frequency, content, and length of messages. Our study results indicated that females were generally more satisfied with the overall program than males. Generally, the relationship between user satisfaction with health services and self-reported gender seems inconclusive. In a systematic review of 39 studies, the majority of the studies (66.7%) showed that there was no significant relationship between the two factors, and the rest were nearly equally divided in terms of favoring either males or females [36].

Self-reported levels of the ability to cope with psychiatric burdens was mostly lower in Text4Hope respondents than in respondents from the Text4Mood study by Agyapong et al [5]. This was true for respondents with depression (56.1% vs 76.7%) and those who experienced loneliness (48.5% vs 57%). However, our results on participants' ability to cope with stress symptoms were consistent with those of Agyapong et al [5]

(77.1% vs 77.2%). These differences could be attributed to the unprecedented COVID-19 pandemic, associated distress, and the strict pandemic-related restrictions (eg, self-isolation and quarantine). These restrictions may be perceived as limitations of personal freedom and activity and may contribute to feelings of loneliness. Similarly, while the perceived improvement in quality of life scores was positive for more than half of our respondents (1057/1714, 61.7%), it was about 14% lower than that of the Agyapong et al study [5]. This may reflect the potentially high negative and multifocal impacts of the COVID-19 pandemic on people's perceived quality of life. In addition, females reported high satisfaction with the Text4Hope program's ability to help them cope with loneliness and depression. This may be in line with the view that depressive symptoms are more frequently experienced by females [37] than males and the fact that people are usually more willing to participate in research that is related to a condition or disease that they have experienced [38]. Text4Hope therefore seems to be a useful support service that helps to ameliorate distressing symptoms in this differentially affected group.

More than 70% of the people in our study reported that the Text4Hope messages were always positive (1336/1732, 77.1%), affirmative (1231/1727, 71.3%), and succinct (1254/1722, 72.8%). About 60% of respondents reported that the messages were always relevant (1052/1753, 61.1%). These results typically came from females, who are usually satisfied with texting services and actively interact with such text messages [19]. Our satisfaction rates were higher than the rates reported by Agyapong et al [5], which ranged from 45.1% to 60%. Similarly, the feeling of being connected to the health care system received higher positive response rates than those in the Agyapong et al study [5] (81% vs 75.2%, respectively). This result may reflect Alberta residents' true need to connect with a health care system during the absence of the regular, conventional care that was provided before the COVID-19 pandemic, given that all of our subscribers were actively seeking help through the texting program.

The number of Text4Hope respondents who reported that they always or often read the text messages was similar to that of the 2016 Agyapong et al study [5] and higher than that of the 2013 Agyapong et al study [39] (84%). Additionally, more than half our subscribers (1125/1716, 65.6%) reported that they always, often, or sometimes returned to the text messages. This is fairly comparable to the Agyapong et al study [5], wherein 33% of respondents reported that they returned to text messages more than once, with no gender differences observed. This is also consistent with the Bendsten & Bendsten study, which reported no differences based on gender in students' satisfaction with a texting service for alcohol use disorder [14]. Consistent with the observations in the study by Agyapong et al [5], the majority of our respondents (1471/1666, 88.3%) reported that they reflected on text messages or took positive actions after reading the text messages, and we believe this could be attributed to the reported positive impact of the program on respondents.

With regard to subscribers' anticipated agreement with the provision of diverse, technology-based medical services, our respondents generally praised the use of these services during

the COVID-19 pandemic and other similar crises. Compared to the other proposed technology-based medical services, our results showed that text messaging was the most highly accepted intervention, with an overall agreement rate of 87.8% (1465/1669). This could be explained by the simple nature of such programs, which is important to the end users who usually own cell phones, and by the short and easy-to-read nature of the daily text messages.

Our study reported slightly lower levels of acceptance for video consultation services for both mental and physical health compared to those for web-based counseling services. This may be attributed to the lack of required physical interaction in video consultation services, as the one-way nature of web-based counseling services is usually more accepted and welcomed by users [40]. However, when therapeutic interaction is required, users may prefer face-to-face services, especially in times of global crises, due to privacy concerns related to therapy in the context of web communication [41]. Additionally, the physical presence of a therapist could play a therapeutic role and promote more interaction, subsequently improving resilience and overall psychological outcomes, especially on a long-term basis [39].

This study has several limitations. For instance, there was a low response rate (9.1%) among the 6-week subscribers, which may have been due to the incentive-free and optional nature of the survey. Thus, the reported levels of satisfaction may have been skewed if there was a systematic difference in the measured features between responders and nonresponders. Notwithstanding the low response rate, our sample size exceeded the 1775 respondents needed to estimate satisfaction rates for the entire subscriber population with a 3% margin of error and 99% confidence. Consequently, our study was sufficiently powered to provide satisfaction rate estimates for the entire population of Text4Hope subscribers. Furthermore, Text4Hope has achieved a higher retention rate than those of other mental health apps that target anxiety, depression, or emotional well-being [18,34]. This high retention rate potentially reflects Text4Hope user satisfaction, which may not be captured through surveys for which completion may be considered time-consuming by some subscribers.

It is also possible that we achieved high satisfaction because people who like technology may have been drawn to the Text4Hope program. Additionally, there is potential for social desirability bias, which may have resulted in respondents reporting higher satisfaction and better perceived benefits from receiving text messages. However, this is unlikely due to the anonymous nature of the survey.

There are several other possible limitations. It is possible that our finding that texting was the most accepted mode of delivery for technology-based health services was biased, as those who liked text messaging were likely to sign up for Text4Hope and therefore participate in the survey. It would have been ideal to include a control group for the comparison of Text4Hope subscribers' and nonsubscribers' anticipated receptivity to technology based medical services. Additionally, although there was a statistically significant gender difference in overall satisfaction between males and females ( $P < .001$ ), the magnitude of the difference was very small and unlikely to be practically meaningful, especially given the imbalance of gender identity subsample sizes. Similarly, our study population was skewed toward females, which is not representative of the population in Alberta or Canada. Finally, respondents' feedback regarding their ability to cope with psychiatric conditions was self-assessed and was not corroborated by clinical assessments.

In conclusion, our results indicate that texting-based programs are acceptable to end users, as high overall satisfaction was reported by subscribers of all gender identities. However, female subscribers reported significantly higher satisfaction scores than male subscribers. Our respondents affirmed the high quality of the text messages by consistently reading and rereading the text messages and providing positive feedback regarding the messages' supportive nature. Text-based mental health support services can be easily deployed during pandemics to support at-risk populations and alleviate the negative mental health impacts that have been well-documented during uncertain times. Based on Text4Hope subscriber feedback, messages from text-based support interventions that have a 160-character limit, are written by health professionals, and are delivered daily can result in high levels of acceptance and satisfaction upon implementation.

---

## Acknowledgments

Support for this study was received from Alberta Health Services and the University of Alberta. This study was supported by grants from the Mental Health Foundation, the Calgary Health Trust, the University Hospital Foundation, the Alberta Children's Hospital Foundation, the Royal Alexandra Hospital Foundation, and the Alberta Cancer Foundation. The funders had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, and approval of the manuscript; or the decision to submit the results for publication.

---

## Authors' Contributions

VIOA conceived and designed the study and the Text4Hope program with MH. RS drafted the initial manuscript with VIOA. AG, WV, and SS participated in data collection. All authors contributed to the study design and revised and approved the final draft of the manuscript.

---

## Conflicts of Interest

None declared.

---

## References

1. Listings of WHO's response to COVID-19. World Health Organization. 2020 Jun 29. URL: <https://www.who.int/news-room/detail/29-06-2020-covidtimeline> [accessed 2020-08-06]
2. Coronavirus disease 2019 (COVID-19) Situation Report – 63. World Health Organization. 2020 Mar 23. URL: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200323-sitrep-63-covid-19.pdf?sfvrsn=b617302d\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200323-sitrep-63-covid-19.pdf?sfvrsn=b617302d_4) [accessed 2020-09-11]
3. Agyapong VIO. Coronavirus disease 2019 pandemic: Health system and community response to a text message (Text4Hope) program supporting mental health in Alberta. *Disaster Med Public Health Prep* 2020 Oct;14(5):e5-e6 [FREE Full text] [doi: [10.1017/dmp.2020.114](https://doi.org/10.1017/dmp.2020.114)] [Medline: [32317038](https://pubmed.ncbi.nlm.nih.gov/32317038/)]
4. Text4Hope: Supporting mental health & wellness in a time of stress & isolation. Alberta Health Services. URL: <https://www.albertahealthservices.ca/topics/Page17019.aspx> [accessed 2020-04-11]
5. Agyapong VIO, Mrklas K, Juhás M, Omeje J, Ohinmaa A, Dursun SM, et al. Cross-sectional survey evaluating Text4Mood: mobile health program to reduce psychological treatment gap in mental healthcare in Alberta through daily supportive text messages. *BMC Psychiatry* 2016 Nov 08;16(1):378. [doi: [10.1186/s12888-016-1104-2](https://doi.org/10.1186/s12888-016-1104-2)] [Medline: [27821096](https://pubmed.ncbi.nlm.nih.gov/27821096/)]
6. Urichuk L, Hrabok M, Hay K, Spurvey P, Sosdjan D, Knox M, et al. Enhancing peer support experience for patients discharged from acute psychiatric care: protocol for a randomised controlled pilot trial. *BMJ Open* 2018 Aug 17;8(8):e022433 [FREE Full text] [doi: [10.1136/bmjopen-2018-022433](https://doi.org/10.1136/bmjopen-2018-022433)] [Medline: [30121611](https://pubmed.ncbi.nlm.nih.gov/30121611/)]
7. Duan L, Zhu G. Psychological interventions for people affected by the COVID-19 epidemic. *Lancet Psychiatry* 2020 Apr;7(4):300-302. [doi: [10.1016/S2215-0366\(20\)30073-0](https://doi.org/10.1016/S2215-0366(20)30073-0)] [Medline: [32085840](https://pubmed.ncbi.nlm.nih.gov/32085840/)]
8. Chen J. Pathogenicity and transmissibility of 2019-nCoV-A quick overview and comparison with other emerging viruses. *Microbes Infect* 2020 Mar;22(2):69-71 [FREE Full text] [doi: [10.1016/j.micinf.2020.01.004](https://doi.org/10.1016/j.micinf.2020.01.004)] [Medline: [32032682](https://pubmed.ncbi.nlm.nih.gov/32032682/)]
9. Public Health Act. Province of Alberta. URL: <https://www.qp.alberta.ca/documents/Acts/P37.pdf> [accessed 2020-04-20]
10. Coronavirus disease 2019 (COVID-19) Situation Report – 65. World Health Organization. 2020 Mar 25. URL: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200325-sitrep-65-covid-19.pdf?sfvrsn=ce13061b\\_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200325-sitrep-65-covid-19.pdf?sfvrsn=ce13061b_2) [accessed 2020-04-20]
11. Agyapong VIO, Juhás M, Ohinmaa A, Omeje J, Mrklas K, Suen VYM, et al. Randomized controlled pilot trial of supportive text messages for patients with depression. *BMC Psychiatry* 2017 Aug 02;17(1):286 [FREE Full text] [doi: [10.1186/s12888-017-1448-2](https://doi.org/10.1186/s12888-017-1448-2)] [Medline: [28768493](https://pubmed.ncbi.nlm.nih.gov/28768493/)]
12. Agyapong VIO, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. *J Affect Disord* 2012 Dec 10;141(2-3):168-176. [doi: [10.1016/j.jad.2012.02.040](https://doi.org/10.1016/j.jad.2012.02.040)] [Medline: [22464008](https://pubmed.ncbi.nlm.nih.gov/22464008/)]
13. Agyapong VIO, Juhás M, Mrklas K, Hrabok M, Omeje J, Gladue I, et al. Randomized controlled pilot trial of supportive text messaging for alcohol use disorder patients. *J Subst Abuse Treat* 2018 Nov;94:74-80. [doi: [10.1016/j.jsat.2018.08.014](https://doi.org/10.1016/j.jsat.2018.08.014)] [Medline: [30243421](https://pubmed.ncbi.nlm.nih.gov/30243421/)]
14. Bendtsen M, Bendtsen P. Feasibility and user perception of a fully automated push-based multiple-session alcohol intervention for university students: randomized controlled trial. *JMIR Mhealth Uhealth* 2014 Jun 23;2(2):e30 [FREE Full text] [doi: [10.2196/mhealth.3233](https://doi.org/10.2196/mhealth.3233)] [Medline: [25098296](https://pubmed.ncbi.nlm.nih.gov/25098296/)]
15. Evans WD, Bihm JW, Szekely D, Nielsen P, Murray E, Abroms L, et al. Initial outcomes from a 4-week follow-up study of the Text4baby program in the military women's population: randomized controlled trial. *J Med Internet Res* 2014 May 20;16(5):e131 [FREE Full text] [doi: [10.2196/jmir.3297](https://doi.org/10.2196/jmir.3297)] [Medline: [24846909](https://pubmed.ncbi.nlm.nih.gov/24846909/)]
16. Abroms LC, Johnson PR, Heminger CL, Van Alstyne JM, Leavitt LE, Schindler-Ruwisch JM, et al. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. *JMIR Mhealth Uhealth* 2015 Jan 23;3(1):e10 [FREE Full text] [doi: [10.2196/mhealth.3846](https://doi.org/10.2196/mhealth.3846)] [Medline: [25650765](https://pubmed.ncbi.nlm.nih.gov/25650765/)]
17. Stamboglis N, Jacobs R. Factors associated with patient satisfaction of community mental health services: A multilevel approach. *Community Ment Health J* 2020 Jan;56(1):50-64 [FREE Full text] [doi: [10.1007/s10597-019-00449-x](https://doi.org/10.1007/s10597-019-00449-x)] [Medline: [31522349](https://pubmed.ncbi.nlm.nih.gov/31522349/)]
18. Ferrell OC, Hartline MD. *Marketing Strategy, Text and Cases*, Sixth Edition. Ohio, USA: Cengage Learning; 2012.
19. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the Patient Health Questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Ment Health* 2015 Mar 24;2(1):e8 [FREE Full text] [doi: [10.2196/mental.3889](https://doi.org/10.2196/mental.3889)] [Medline: [26543914](https://pubmed.ncbi.nlm.nih.gov/26543914/)]
20. 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](https://doi.org/10.2196/jmir.7740)] [Medline: [28838887](https://pubmed.ncbi.nlm.nih.gov/28838887/)]
21. Ben-Zeev D, Kaiser SM, Krzos I. Remote "hovering" with individuals with psychotic disorders and substance use: feasibility, engagement, and therapeutic alliance with a text-messaging mobile interventionist. *J Dual Diagn* 2014;10(4):197-203 [FREE Full text] [doi: [10.1080/15504263.2014.962336](https://doi.org/10.1080/15504263.2014.962336)] [Medline: [25391277](https://pubmed.ncbi.nlm.nih.gov/25391277/)]
22. Mohr DC, Ho J, Duffecy J, Reifler D, Sokol L, Burns MN, et al. Effect of telephone-administered vs face-to-face cognitive behavioral therapy on adherence to therapy and depression outcomes among primary care patients: a randomized trial. *JAMA* 2012 Jun 06;307(21):2278-2285 [FREE Full text] [doi: [10.1001/jama.2012.5588](https://doi.org/10.1001/jama.2012.5588)] [Medline: [22706833](https://pubmed.ncbi.nlm.nih.gov/22706833/)]

23. Garcia C, Hardeman RR, Kwon G, Lando-King E, Zhang L, Genis T, et al. Teenagers and texting: use of a youth ecological momentary assessment system in trajectory health research with latina adolescents. *JMIR Mhealth Uhealth* 2014 Jan 24;2(1):e3 [FREE Full text] [doi: [10.2196/mhealth.2576](https://doi.org/10.2196/mhealth.2576)] [Medline: [25098355](https://pubmed.ncbi.nlm.nih.gov/25098355/)]
24. Irvine L, Falconer DW, Jones C, Ricketts IW, Williams B, Crombie IK. Can text messages reach the parts other process measures cannot reach: an evaluation of a behavior change intervention delivered by mobile phone? *PLoS One* 2012;7(12):e52621. [doi: [10.1371/journal.pone.0052621](https://doi.org/10.1371/journal.pone.0052621)] [Medline: [23300726](https://pubmed.ncbi.nlm.nih.gov/23300726/)]
25. Population projections. Alberta and Census Divisions, 2020 – 2046. URL: <https://open.alberta.ca/dataset/90a09f08-c52c-43bd-b48a-fda5187273b9/resource/bb7c6ef6-ade5-4def-ae55-ef1fd5d4e563/download/2020-2046-alberta-population-projections.pdf> [accessed 2020-11-14]
26. Alberta population. World Population Review. URL: <https://worldpopulationreview.com/canadian-provinces/alberta-population> [accessed 2020-12-21]
27. Education in Canada: Key results from the 2016 Census. Statistics Canada. 2017 Nov 29. URL: <https://www150.statcan.gc.ca/n1/daily-quotidien/171129/dq171129a-eng.htm> [accessed 2020-12-21]
28. Gross domestic product. Government of Alberta. URL: <https://economicdashboard.alberta.ca/GrossDomesticProduct> [accessed 2020-11-13]
29. Agyapong VIO, Hrabok M, Vuong W, Gusnowski A, Shalaby R, Mrklas K, et al. Closing the Psychological Treatment Gap During the COVID-19 Pandemic With a Supportive Text Messaging Program: Protocol for Implementation and Evaluation. *JMIR Res Protoc* 2020 Jun 22;9(6):e19292 [FREE Full text] [doi: [10.2196/19292](https://doi.org/10.2196/19292)]
30. Nwachukwu I, Nkire N, Shalaby R, Hrabok M, Vuong W, Gusnowski A, et al. COVID-19 pandemic: Age-related differences in measures of stress, anxiety and depression in Canada. *Int J Environ Res Public Health* 2020 Sep 01;17(17):6366 [FREE Full text] [doi: [10.3390/ijerph17176366](https://doi.org/10.3390/ijerph17176366)] [Medline: [32882922](https://pubmed.ncbi.nlm.nih.gov/32882922/)]
31. Abba-Aji A, Li D, Hrabok M, Shalaby R, Gusnowski A, Vuong W, et al. COVID-19 pandemic and mental health: Prevalence and correlates of new-onset obsessive-compulsive symptoms in a Canadian province. *Int J Environ Res Public Health* 2020 Sep 24;17(19):6986 [FREE Full text] [doi: [10.3390/ijerph17196986](https://doi.org/10.3390/ijerph17196986)] [Medline: [32987764](https://pubmed.ncbi.nlm.nih.gov/32987764/)]
32. Release notes - IBM® SPSS® Statistics 26.0. IBM Corporation. URL: <https://www.ibm.com/support/pages/release-notes-ibm%C2%AE-spss%C2%AE-statistics-260> [accessed 2020-03-30]
33. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. *Am J Prev Med* 2009 Feb;36(2):165-173. [doi: [10.1016/j.amepre.2008.09.040](https://doi.org/10.1016/j.amepre.2008.09.040)] [Medline: [19135907](https://pubmed.ncbi.nlm.nih.gov/19135907/)]
34. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *J Med Internet Res* 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: [10.2196/14567](https://doi.org/10.2196/14567)] [Medline: [31573916](https://pubmed.ncbi.nlm.nih.gov/31573916/)]
35. Berrouguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): A systematic review of mobile phone and web-based text messaging in mental health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [FREE Full text] [doi: [10.2196/jmir.5066](https://doi.org/10.2196/jmir.5066)] [Medline: [27287668](https://pubmed.ncbi.nlm.nih.gov/27287668/)]
36. Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002;6(32):1-244. [doi: [10.3310/hta6320](https://doi.org/10.3310/hta6320)] [Medline: [12925269](https://pubmed.ncbi.nlm.nih.gov/12925269/)]
37. Kang HJ, Park Y, Yoo KH, Kim KT, Kim ES, Kim JW, et al. Sex differences in the genetic architecture of depression. *Sci Rep* 2020 Jun 18;10(1):9927 [FREE Full text] [doi: [10.1038/s41598-020-66672-9](https://doi.org/10.1038/s41598-020-66672-9)] [Medline: [32555505](https://pubmed.ncbi.nlm.nih.gov/32555505/)]
38. Glass DC, Kelsall HL, Slegers C, Forbes AB, Loff B, Zion D, et al. A telephone survey of factors affecting willingness to participate in health research surveys. *BMC Public Health* 2015 Oct 05;15:1017 [FREE Full text] [doi: [10.1186/s12889-015-2350-9](https://doi.org/10.1186/s12889-015-2350-9)] [Medline: [26438148](https://pubmed.ncbi.nlm.nih.gov/26438148/)]
39. Agyapong VIO, Milnes J, McLoughlin DM, Farren CK. Perception of patients with alcohol use disorder and comorbid depression about the usefulness of supportive text messages. *Technol Health Care* 2013;21(1):31-39. [doi: [10.3233/THC-120707](https://doi.org/10.3233/THC-120707)] [Medline: [23358057](https://pubmed.ncbi.nlm.nih.gov/23358057/)]
40. Brenes GA, Danhauer SC, Lyles MF, Hogan PE, Miller ME. Telephone-delivered cognitive behavioral therapy and telephone-delivered nondirective supportive therapy for rural older adults with generalized anxiety disorder: A randomized clinical trial. *JAMA Psychiatry* 2015 Oct;72(10):1012-1020 [FREE Full text] [doi: [10.1001/jamapsychiatry.2015.1154](https://doi.org/10.1001/jamapsychiatry.2015.1154)] [Medline: [26244854](https://pubmed.ncbi.nlm.nih.gov/26244854/)]
41. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: Systematic literature review. *J Med Internet Res* 2020 Feb 20;22(2):e16407 [FREE Full text] [doi: [10.2196/16407](https://doi.org/10.2196/16407)] [Medline: [32130131](https://pubmed.ncbi.nlm.nih.gov/32130131/)]

## Abbreviations

**ANOVA:** analysis of variance

**MERS-CoV:** Middle East respiratory syndrome coronavirus

**SARS-CoV:** severe acute respiratory syndrome coronavirus



*Edited by L Buis; submitted 08.09.20; peer-reviewed by E Kleiman, E Godfrey, P Jimenez, R Milhausen; comments to author 03.11.20; revised version received 23.11.20; accepted 08.03.21; published 15.04.21.*

*Please cite as:*

*Shalaby R, Vuong W, Hrabok M, Gusnowski A, Mrklas K, Li D, Snaterse M, Surood S, Cao B, Li XM, Greiner R, Greenshaw AJ, Agyapong VIO*

*Gender Differences in Satisfaction With a Text Messaging Program (Text4Hope) and Anticipated Receptivity to Technology-Based Health Support During the COVID-19 Pandemic: Cross-sectional Survey Study*

*JMIR Mhealth Uhealth 2021;9(4):e24184*

*URL: <https://mhealth.jmir.org/2021/4/e24184>*

*doi: [10.2196/24184](https://doi.org/10.2196/24184)*

*PMID: [33750738](https://pubmed.ncbi.nlm.nih.gov/33750738/)*

©Reham Shalaby, Wesley Vuong, Marianne Hrabok, April Gusnowski, Kelly Mrklas, Daniel Li, Mark Snaterse, Shireen Surood, Bo Cao, Xin-Min Li, Russell Greiner, Andrew James Greenshaw, Vincent Israel Opoku Agyapong. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 15.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Designing an App to Overcome Language Barriers in the Delivery of Emergency Medical Services: Participatory Development Process

Eva Maria Noack<sup>1</sup>, Dr; Jennifer Schulze<sup>1</sup>; Frank Müller<sup>1</sup>, Dr med

Department of General Practice, University Medical Center Göttingen, Göttingen, Germany

**Corresponding Author:**

Eva Maria Noack, Dr

Department of General Practice

University Medical Center Göttingen

Humboldtallee 38

Göttingen, 37073

Germany

Phone: 49 55139 68193

Email: [evamaria.noack@med.uni-goettingen.de](mailto:evamaria.noack@med.uni-goettingen.de)

## Abstract

**Background:** In emergencies, language barriers may have dangerous consequences for the patients. There have been some technical approaches to overcome language barriers in medical care but not yet in the prehospital emergency care setting. The use of digital technologies in health care is expanding rapidly. Involving end users at all stages of the development process may help to ensure such technologies are usable and can be implemented.

**Objective:** We aimed to develop a digital communication tool that addresses paramedic needs in the specific circumstances of prehospital emergency care and helps paramedics to overcome language barriers when providing care to foreign-language patients.

**Methods:** We actively engaged paramedics and software designers in an action-oriented, participatory, iterative development process, which included field observations, workshops, background conversations, questionnaires on rescue missions, studying the literature, and preliminary testing in the field.

**Results:** With input from paramedics, we created an app with 600 fixed phrases supporting 18 languages. The app includes medical history-taking questions, phrases asking for consent, and phrases providing specific additional information. Children as patients, as well as their carers and other third parties, can be addressed with appropriate wording. All phrases can be played back audibly or displayed as text. The comprehensive content is grouped into categories and adapted to diverse scenarios, which makes the tool rapidly usable. The app includes a function to document patient responses and the conversation history. For evaluation in a clinical study, the app is run on a smartphone with extra speakers to be of use in noisy environments. The use of prototypes proved valuable to verify that the content, structure, and functions discussed in theory were of value and genuinely needed in practice and that the various device control elements were intuitive.

**Conclusions:** The nature of the paramedic work environment places specific demands on the communication options used and need for such devices. The active involvement of paramedics in the development process allowed us to understand and subsequently consider their experience-based knowledge. Software designers could understand the paramedics' work environment and consider respective needs in the menu navigation and design principles of the app. We argue that the development of any medical software product should actively involve both end users and developers in all phases of the development process. Providing the users with the opportunity to influence technology development ensures that the result is closer to their needs, which can be seen as crucial for successful implementation and sustainable use.

**Trial Registration:** German Clinical Trials Register DRKS00016719; [https://www.drks.de/drks\\_web/navigate.do?navigationId=trial.HTML&TRIAL\\_ID=DRKS00016719](https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00016719)

**International Registered Report Identifier (IRRID):** RR2-10.1186/s12913-020-05098-5

(*JMIR Mhealth Uhealth* 2021;9(4):e21586) doi:[10.2196/21586](https://doi.org/10.2196/21586)

## KEYWORDS

paramedic; interpreter; medical translation; application software; app; digital communication tool; foreign-language patients; language barrier; participatory design; prehospital emergency care; emergency medical service

## Introduction

For medical treatment, it is essential that patients and medical staff can communicate with each other. Language barriers can have a negative impact on the quality of care and patient outcomes [1,2]. Particularly, in out-of-hospital emergency situations that require rapid assessments and decision making and frequently necessitate initial treatment on the scene, language barriers can have dangerous consequences for patients [3,4]. In a more rural area of Germany, 2.2% of the rescue service operations involved patients with a language barrier [5]. This number is expected to be significantly higher in urban areas. In emergencies, it is therefore crucial for the paramedic staff to understand the patients' acute complaints, pre-existing conditions, allergies, or drug treatments. However, at the emergency scene, accurate translation by professional interpreters is rarely available. Therefore, paramedics mostly rely on nonverbal communication, use a third language, or rely on lay interpreters [6,7]. Translation accuracy cannot be guaranteed when making use of interpretation support from bystanders, and confidentiality problems arise [1,2].

There are several technical solutions to improve patient-staff communication in multilingual clinical settings. In hospital and primary care, video or telephone interpreting services may be used [8-10]. For time-critical emergencies on potentially dangerous sites, these are often not available and would require a reliable network coverage, which is rarely guaranteed in rural areas in Germany. The quality of machine translation such as Google Translate technology is improving, but it is still not considered accurate enough for actual deployment in health settings [11,12], and similarly, ad-hoc translators require a permanent connection to the internet. In addition to reliability issues, data confidentiality problems are unsolved.

Communication tools to overcome language barriers in health care must be developed and suitably adapted to the specific working environment and health care setting. Examples include a tablet-based app for medical history taking for refugees and asylum seekers in a German transit camp, which was developed and piloted in a family doctor surgery setting [13,14]; a speech-enabled, fixed-phrase translator was found to be a good alternative to collecting information if interpreters are not available in an emergency room setting [15]. However, these tools do not suit prehospital emergency care situations, which are frequently complex, volatile, and rushed. A tool for rescue operations therefore requires not only adapted content but also a different communication approach.

In the project "DICTUM rescue" (Digital cross-cultural interpreter-tool in medical consultations for refugees and migrants), we aim to develop a digital communication tool that helps paramedics to overcome language barriers when providing care to foreign-language patients in the prehospital emergency care setting.

Digital technologies have the potential to simplify work processes and improve productivity. However, if such tools do not meet user needs or are inconvenient to use, they will be used reluctantly or not at all. Involving end users may make digital innovations more suitable for their use, which will consequently improve technology uptake. While there are examples of patients, medical professionals, and other stakeholders being successfully involved in the design of medical technology [16-18], they rarely take part in the development process.

By involving paramedics and software designers in the development, we aimed to develop an app that meets the needs of paramedics when providing care to foreign-language patients and accommodates the specific circumstances of rescue operations. In this paper, we focus on the app development process and paramedics' active involvement in this process.

The app will be evaluated in an interventional trial (No. DRKS00016719) [19].

## Methods

### Action-Oriented Participatory Approach: Designing for, With, and by the Users

We used an action-oriented participatory approach, actively involving paramedics from 4 emergency medical service stations in Lower Saxony, Germany. Action-oriented approaches are devoted to defining problems and finding respective solutions in real-world situations together with the people who experience them [20,21]. In the development and implementation of health technologies, participatory designs have been used and perceived as helpful [22,23]. Our approach combined aspects that can be found in "classical" participatory research as well as in participatory design. While the first usually operates with clear, predefined research questions and endpoints, the latter is described as a process-based and rather open development, where an object, tool, or service is designed [24]. The idea behind our approach was: If this app is to succeed in improving paramedic communication with foreign-language patients, it has to meet the specific needs of paramedics and accommodate the particular setting of emergency medical services. Software designers (aidminutes GmbH) were similarly engaged in the participatory approach from the beginning so they could understand paramedic expectations and requirements for rescue missions. The research team consisted of 2 medical scientists and a study nurse, a trained paramedic. The development process and documentation should strictly follow scientific criteria and include external evidence, for example from literature, and meet ethical standards of qualitative research. For the realization of the project, we obtained ethical approval from the research ethics board of the University Medical Center Göttingen (Ethics Approval No. 9/9/18).

We used several techniques to investigate specific aspects for content and technical development (Table 1). An extensive literature search of relevant textbooks and guidelines

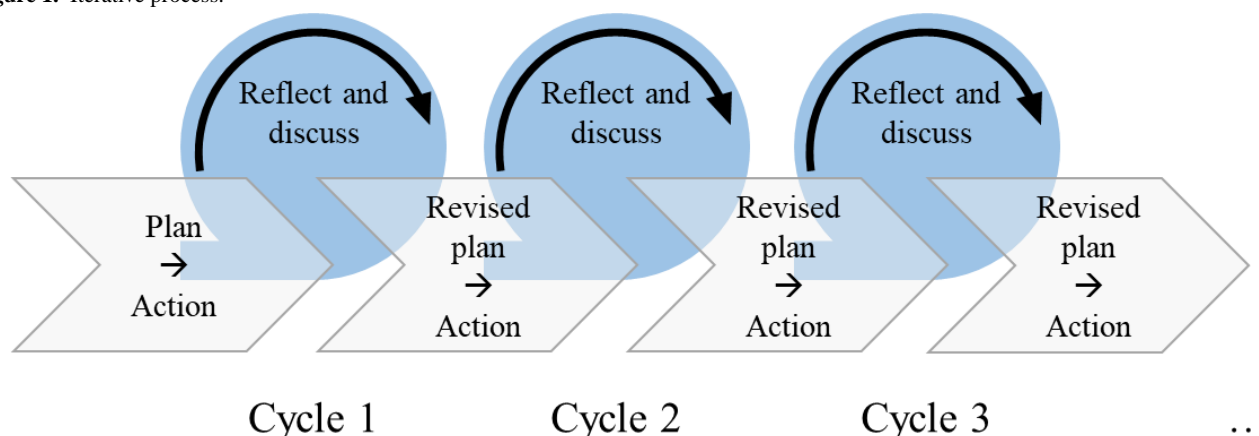
accompanied the development process. Data were collected using records, field notes, and minutes.

**Table 1.** Overview of the techniques used for app design.

Timeframe	Activity	Participants	Aim	Data collection and analysis
Dec 2018 - Mar 2019	Field observations (“internship”)	Software designers	Experience paramedics’ daily work, gain knowledge regarding the nature of the role and its specific challenges.	Observations as to the requirements and possible challenges for the app were discussed within the team and with paramedics.
Feb 2019 - Sep 2019	Workshops	Paramedics from the rescue stations and software designers	Different foci and formats: Discuss and define content and structure; simulate cases in role play to evaluate content and structure using paper prototypes; identify useful functions; check hardware options.	Paramedics discussed content and structure and reflected on the role plays. Findings (minutes of discussions and suggestions, observations of simulations) were summarized by the research team, discussed, checked with literature and guidelines, and included in new versions. Findings as to functionality were discussed with software designers.
Jan 2019 - Dec 2019	Background conversations	Paramedics and relevant management (heads and deputies of the rescue stations, the district’s rescue service and fire department)	Provide a quick opinion and preliminary appraisal of a feature, wording or content; test app prototypes and assess perception; keep the staff informed about the progress of the project; give paramedics the ongoing opportunity to share their experience and thoughts with research team; plan implementation in daily routine.	Discussions were held within the research team and with software designers.
Feb 2019 - (ongoing)	Questionnaires	Paramedics	Gain information on real emergencies with foreign-language patients; enable paramedics to promptly describe their respective experience and impressions; detect which languages have to be supported by the app.	Descriptive statistics (scaled questions), frequency analysis (languages), content analyses (free texts); challenges raised could be reflected on the workshops.
Sep 2019 - Oct 2019	Case simulations	Students in the School for Paramedics	External perception of the developed app; test and evaluate the app in a mock real-life setting; finalize the app.	Observation of the simulations by the research team and software designers; students provide feedback on user experience.
Dec 2019	Final (“prerelease”) testing	Two paramedics from each rescue station	Test and evaluate the app in a real-life setting.	Direct feedback to developers and researchers.

The research and development process moved in iterative cycles (Figure 1). At each stage, findings from the previous cycle were reflected on and formed the basis for planning the next phase. For example, intermediate results concerning the communication content (which phrases should be integrated?) and its structure (how should these phrases be grouped?) were discussed and tested in role play. As new components arose, the structure was

revised, the proposed content was adjusted to reflect up-to-date literature and guidelines, features were rethought with technical developers, and the features were set up for discussion and tested again. The number of iterations could not be predetermined beforehand as interesting outputs could not be anticipated and could arise at any stage.

**Figure 1.** Iterative process.

### Workshops and Background Conversations

We invited the rescue station staff to participate in various workshops. These had different formats, including discussions and role plays, adapted to the shifting foci, for example the app's content, structure, functions, appearance, device options, and implementation in daily routine. To test the app content and structure, we simulated rescue missions with foreign-language patients using paper prototypes presented in booklet format (see [Multimedia Appendix 1](#)).

We conducted 8 workshops with 4-8 participants each, which lasted 2-3 hours. We tried to ensure variation in participant work experience, age, and gender. In total, 47 paramedics participated in the workshops.

To obtain feedback on usability and appearance, we presented design mock-ups, click dummies, and finally alpha and beta versions of the app. We also encouraged workshop participants to verbalize their opinions and to "think aloud" while carrying out app-related tasks. This helped us to understand how paramedics experienced and assessed the user interface and to test whether they used control elements and found the desired phrases intuitively.

### Questionnaires on Rescue Missions With Foreign-Language Patients

During the app development phase and as part of the clinical trial, paramedics from the participating rescue stations completed questionnaires after they had provided care to foreign-language patients. These questionnaires provided us with information on how paramedics had experienced the communication with patients and how they had dealt with obstacles. This information complemented the findings from the other activities and could be included in the discussions with paramedics. The questionnaires also helped us to decide which languages should be supported by the app to meet local demand.

### Case Training Simulating Emergency Situations for External Testing

To make sure that the app could be used intuitively by paramedics who had not been involved in the development

process, we tested early versions of the app with third-year and final-year paramedic students. At the school for paramedics, case simulations for training purposes are comprised of a complete set of emergency equipment for a rescue mission. During a simulation, the instructor can remotely control the devices, for example changing a patient's vital parameters and thereby creating a more realistic working environment. We ran 16 case simulations including internal medicine, traumatology, gynecology, and pediatrics in 2 student groups consisting of 5 and 6 students, respectively. Scenarios were prepared by the instructor who also did not know the app. In each simulation, a team of 2 students treated a Dari-speaking patient. Afterwards, the simulations were discussed with the instructors and the students observing the simulation. The researchers and software designers additionally observed the interactions.

### Prerelease Testing of the App for Final Refinements and Implementation

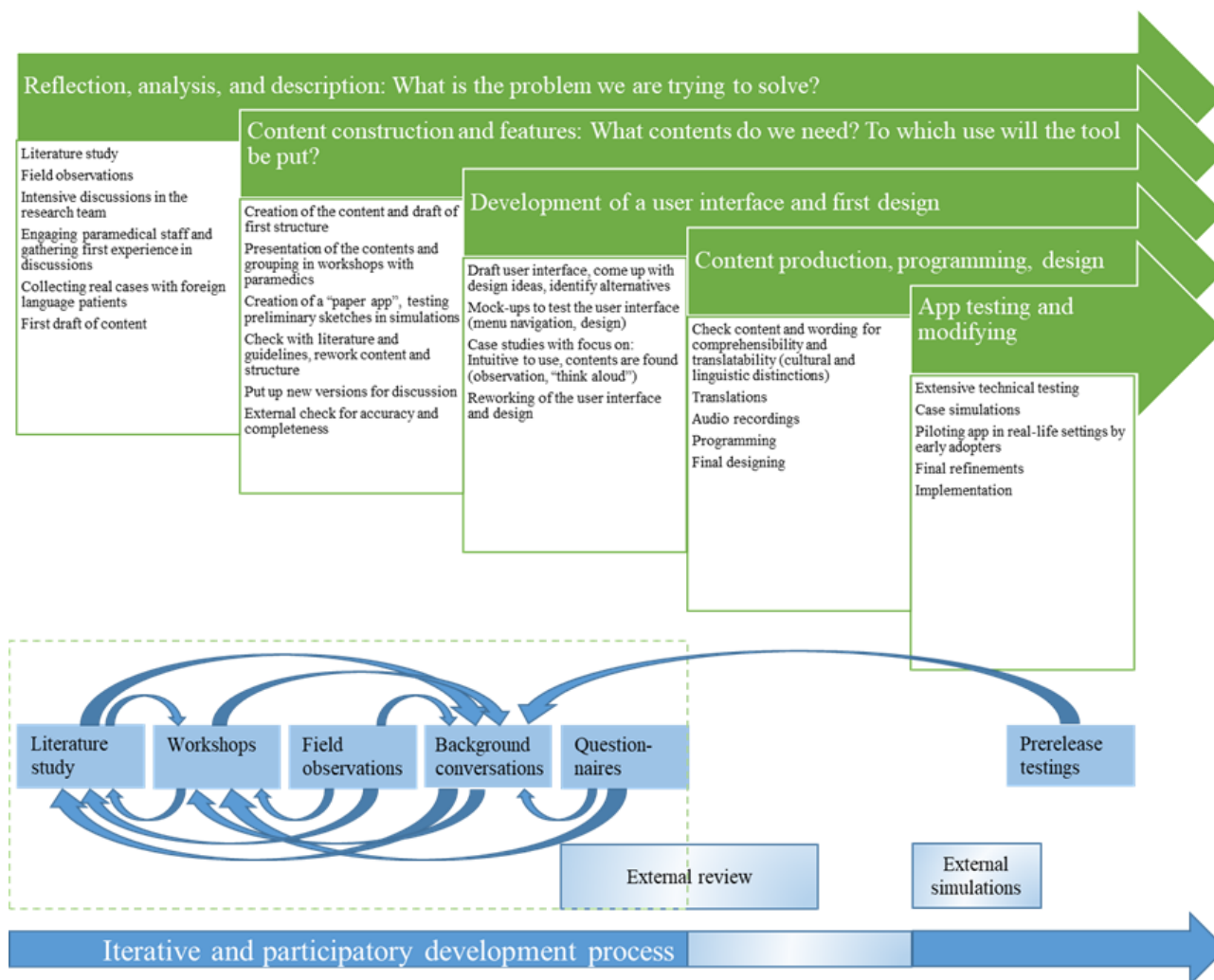
At each rescue station, 2 paramedics tested the app and were asked to provide feedback to the technical developers and researchers for final refinements before implementation. The aim in this phase was to observe performance and the stability of the app in real rescue operations. We established a private messenger channel for direct feedback.

## Results

### Developing in Loops

We started with initial reflections on the field and on app principles. Subsequently, we collected and drafted questions and other phrases that reflected communication situations in emergency medical services. Concurrently, we set up a structure that allowed rapid and flexible medical history taking, adapted to a diverse array of scenarios. We defined the necessary functions that had to be considered in the design of the user interface, explored which devices best suited rescue missions, discussed how to implement the app in daily routine, and established which languages were needed to meet the local population. Finally, we conducted external app testing sessions. [Figure 2](#) gives an oversight of the process of the app development and shows the main feedback loops.

Figure 2. App development process.



**Initial Reflections**

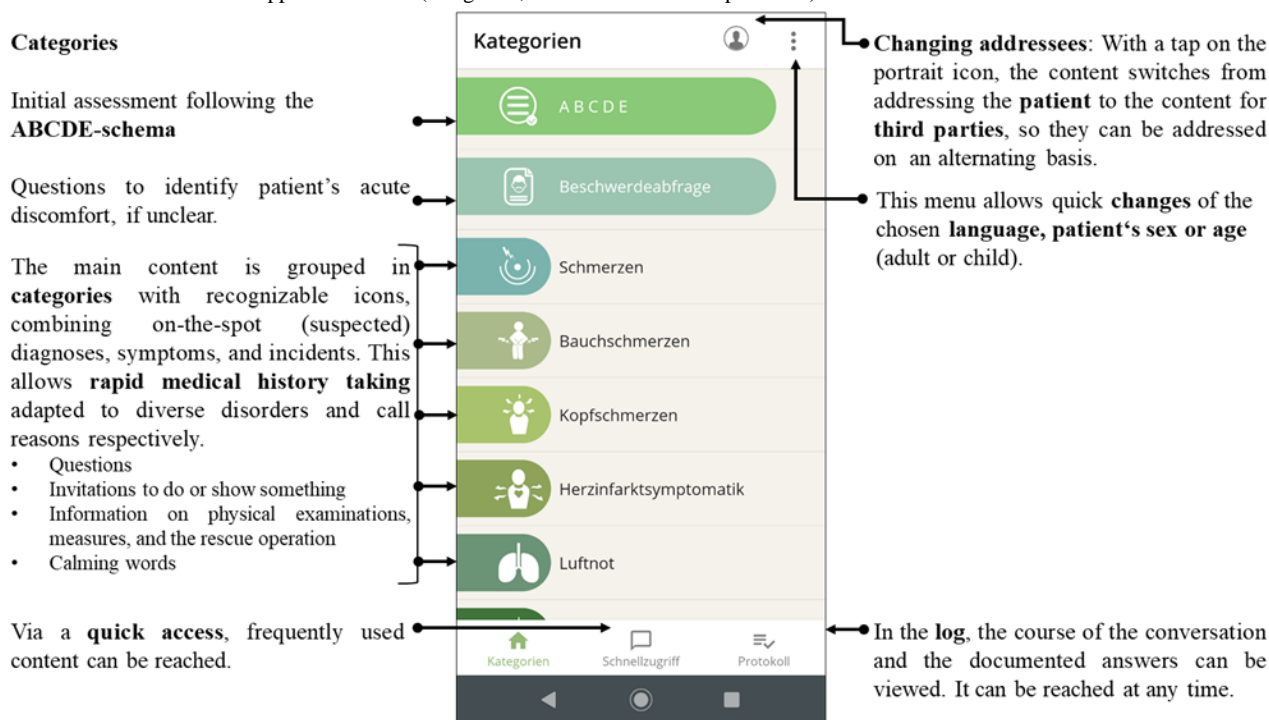
An app that helps overcome language barriers in prehospital emergency situations needs to be reliable, rapid-to-use, and usable offline and has to be run on a handy and sturdy device. The app operator should be the paramedic, as patients might not be able to use a digital device or maintain good hygiene.

To enable paramedic-patient communication that resembles as closely as possible a normal verbal interaction and to include illiterate and myopic patients, all phrases needed to be provided as audible messages. For app use in noisy surroundings, with hearing impaired patients, or to ensure confidential matters would be addressed sensitively and discreetly, phrases should also be displayed in text on the device.

Speech recognition is not yet reliable enough for clinical deployment, especially in situations where the environment is noisy, several speakers are involved, and colloquial expressions or dialects are being used. Therefore, all questions have to be posed in a way that patients can answer them nonverbally; that is, questions have to be closed-ended or contain a request to show something. Patients have to feel confident on how to answer, and the (nonverbal) answers have to be understood unambiguously by the paramedics. If it is not clear which language a patient understands, the patient should be able to select the language she or he prefers. In addition, it should be possible to check whether the patient understands the selected language.

The solution for our study consists of a smartphone-based app. The app’s main functions are outlined in Figure 3 and described in the following sections.

Figure 3. Main functions of the app and structure (categories; see text for further explanation).



### Content

The aim of the first set of activities was to identify the key issues among paramedics in the participating emergency medical stations in the context of rescue missions with foreign-language patients and to determine the communication requirements in such missions.

Paramedics ask questions but they also give information and ask for consent, for example, when they perform measures or (drug) therapy on the patient. Communication also takes place with third parties, for example relatives or bystanders, and patients and third parties are often addressed on an alternating basis.

Pediatric treatment in emergency scenarios requires different content and wording. We therefore devised a child mode that covers children-specific emergencies and provides phrases that can be understood by children. Moreover, the content is tailored to the patient's gender.

These requirements led to an unexpectedly high number of phrases needed, which was a challenge for structuring the content and for the user interface design.

Based on these initial insights, a first draft of questions and phrases was set up for discussion among the paramedics, which included determination of those questions that were indispensable for differential diagnoses, further treatment, the next steps, or for paramedics' (legal) protection and how to reword questions so that they would be closed-ended questions, such as those that asked a number, time, or period of symptom

onset. In this instance, paramedics reflected on the essential information they needed to gather. For example, open questions such as "from what height did you/did the child fall from?" were changed to closed questions "did you fall from higher than 3 meters?" and "can you show me where the child fell from?"

In the workshops, we also worked on phrases for situations that paramedics perceived as challenging, for example dealing with aggressive or intoxicated people, cases of suspected domestic violence, an attempted or actual suicide, or when speaking to family members regarding the fatality of a relative.

Role play helped to establish whether the phrases covered all the important conceivable scenarios and identify unnecessary or unclear content.

Role plays revealed that nonverbal communication also plays an important role in encounters between paramedics and foreign-language patients. The observed nonverbal interaction was discussed with paramedics. As nonverbal communication may be culturally imprinted, the app also contains phrases for situations that one might think can be solved nonverbally. External medical professionals reviewed all phrases as to their accuracy. Before translation, the research team indicated potential ambiguities to prevent misleading translations. Also, all phrases were looked over by a cultural scientist to ensure comprehensibility and translatability. Some phrases though were reworded into simpler language or according to personal preference. In several cases, this changed the original intended meaning in a way that the phrase no longer fitted with (all)

intended situations or did not reflect what was discussed with paramedics, resulting in the need for retranslation.

Subsequently, professional interpreters experienced in the medical field translated and audio-recorded the phrases. Translators were advised to use the most common and comprehensive wording. The results were then reviewed and partly retranslated to ensure high quality of the translations.

The app contains 600 standard phrases, though the tailoring of these for adult patients, pediatric patients, and specific third parties as well as to accommodate the patient's gender, which is needed in many languages, meant that up to 1200 phrases were produced in Arabic, for example, and overall, 16,000 audio-files were generated in total (the app is available for Android and iOS, see [Multimedia Appendix 2](#)).

### Structure

We faced the challenge of grouping the comprehensive content in a way that all phrases could be easily found. Different grouping options, for example according to organ systems, body regions, or potential patient outcomes, were discussed and tested. We also considered fixed series of phrases for certain recurring situations, such as a suspected acute coronary syndrome. A series approach was rejected since the content and course conversations can vary considerably from case to case.

The final adopted categorization approach arose from consideration of a combination of symptoms, on-the-spot (suspected) diagnoses, and incidents, sorted by probability of occurrence [25,26]. There are separate categories for physical examination, informative and reassuring sentences, questions concerning drugs, intolerances, preexisting conditions, and patient documents. Phrases that are necessary for the primary survey following the ABCDE (Airway, Breathing, Circulation, Disability, Exposure) approach form the first group. This systematic approach is recommended by the European Resuscitation Council, is a widely accepted standard of care for the immediate assessment and treatment of life-threatening clinical problems [27], and is applicable in all clinical emergencies.

The second group assembles questions that help uncover the reasons for seeking medical emergency care if not obvious or not known to the paramedics. Phrases may appear in more than one category, and categories are linked with each other for quick navigation. Within categories, the content is clustered according to the paramedic approach of structuring a rescue mission.

### Functions and Navigation

In the second phase of the app development process, the focus shifted toward the process of designing a user interface that reflected the functions that paramedics deemed helpful and allowed easy, rapid, and flexible navigation. In the following sections, we describe the additional functionality requested by the paramedics and how we implemented it.

#### *Audio Playback and Additional Text Display*

All phrases can be both played back audibly or displayed as text. The text is displayed in horizontal format, which facilitates reading for the patient. For languages that do not use the Latin alphabet, the correct reading direction is indicated by an arrow.

#### *Log to Document Patient Responses*

Paramedics indicated that they would like to document the patient responses and to review them at any time. Two possible log views were identified as the most convenient. First, an overview of the documented answers (“yes,” “no,” or “unclear” for closed-ended questions; localization of pain or an injury marked on a figure) is provided using the SAMPLE history scheme (questions for a secondary assessment used in prehospital emergency care including Symptoms, Allergies, Medications, Past medical history, Last oral intake, Events prior to incident) [28]. This function additionally allows paramedics to quickly identify if certain information has not yet been obtained, which therefore facilitates a complete assessment. Second and alternatively, a “chat view” shows the complete course of conversations, including information that was provided.

#### *Quick Access Menu*

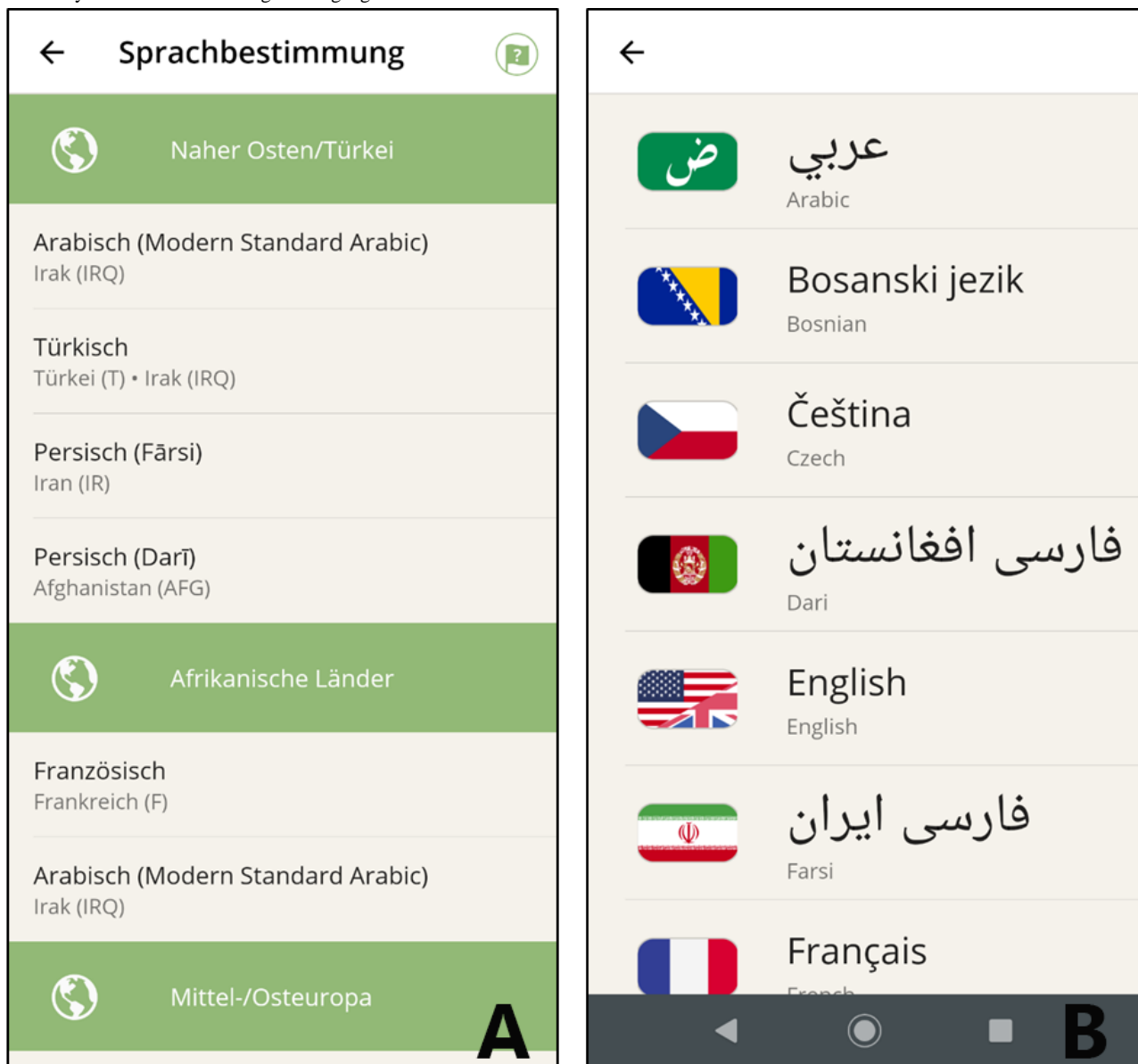
Phrases that are expected to be used frequently are grouped, and these can be accessed by selecting a steady button that serves as a “quick access menu.” We have not implemented the possibility to customize this function because the app is to be installed on a jointly used device on the rescue vehicles.

#### *Navigation*

When starting the app, paramedics have to choose age (adult or child) and gender of the person in need of help. Based on this selection, the app automatically adopts age- and sex-specific content and wording to address children and their health problems, for example. Paramedics then select the language the patient (presumably) speaks from a geographically sorted directory, or the patient selects a language they understand from a flag button list ([Figure 4](#)).



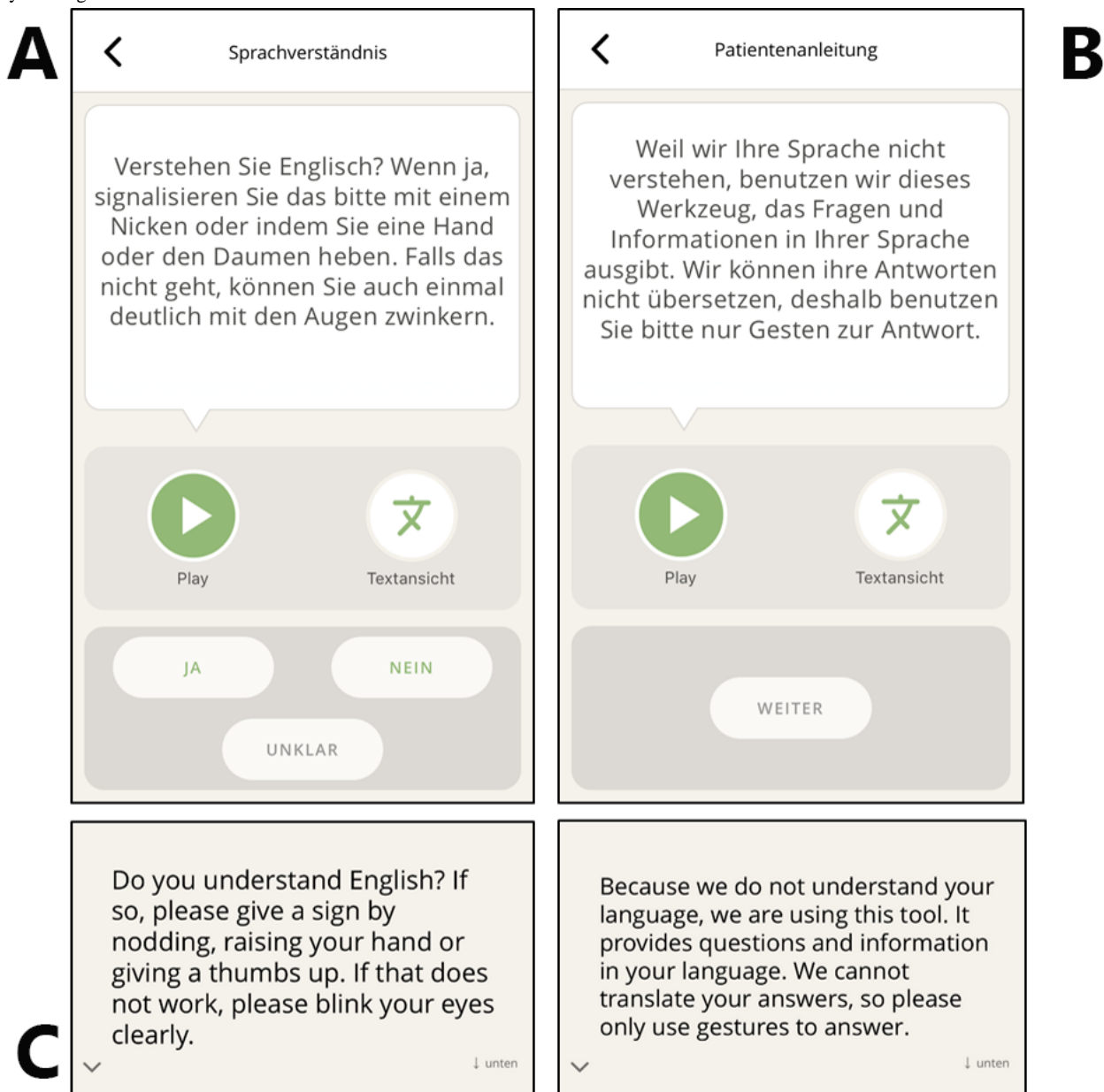
**Figure 4.** Paramedics select the language from (A) a geographically sorted directory or (B) have the patient choose from a flag button list that is sorted alphabetically and labeled in the original language.



Other options, such as trying to ask the patient to show his or her home country on a map, were discarded as some patients may not be able to read a map, while others might be unsure whether the paramedics are asking for their country of birth or where they have travelled from. Additionally, in many countries, more languages and dialects are spoken and cannot be identified with certainty using maps.

A patient's comprehension of the language selected may be checked by asking the patient to give a sign. The used phrase considers that affirmative and rejecting gestures are not universal across cultures and languages. Then, the use of the app is briefly explained to the patients (Figure 5). If required, the language check and app explanations may be skipped to reach the category list more quickly.

**Figure 5.** (A) To confirm that the patient understands the chosen language, the app may ask “Do you understand [chosen language]?” and asks the patient to sign “yes” if so. (B) The app provides a short explanation of the tool’s function to prevent misunderstandings. (C) The respective text is displayed using the “Textansicht” button.



The app navigation takes into account the variety and complexity of rescue missions, for example by allowing the user to switch quickly between addressees (ie, patient or accompanying companions) to reach the patient response log at any time or to change directly to a related category without navigating through category lists again.

In early versions of the app, paramedics experienced difficulties in locating the function to change the addressees and the output language. To address this, icons were redesigned and more appropriately arranged.

### Device Requirements

Emergency medicine places high demands on the used devices: They have to be resistant to falls, dust, water splashes, and extreme temperatures as well as easy to disinfect and portable so that they are easily carried in rescue missions. They need a long-lasting battery, the displayed content must be legible in

sunlight, and the audio output has to be loud enough for noisy situations or to address hearing impaired patients. Different smartphones and tablets and accessories were discussed and tested with the paramedics and the software designers.

For the study, we decided to use a Motorola Play Z2 Android smartphone with an attachable extra speaker, dedicated battery, a sound output of up to 84 decibels, a turbo power charger, a scratch-and crack-resistant glass cover, and a shock-absorbing bumper.

### Languages and Dialects

The selection of languages that needed to be supported by the study version of the app was based on the questionnaires to paramedics regarding recent rescue missions with foreign-language patients, demographic statistics, and cross-border travel data. For this study, the app supports 18 languages and dialects: Arabic, Bosnian, Croatian, Czech, Dari

(Persian), English, Farsi (Persian), French, German, Italian, Kurdish-Sorani, Lithuanian, Pashto (Afghani), Polish, Russian, Serbian, Spanish, and Turkish.

### Experience Obtained From the Development Process

We used a variety of techniques to explore obstacles faced by paramedics in rescue missions with foreign-language patients and to devise an app to tackle these. Our approach allowed both researchers and software designers to understand the paramedics' daily work, perspectives, priorities, and experienced problems as well as their needs and expectations towards an intervention aimed at overcoming language barriers. Paramedics discussed phrases, app structure, and scenarios and tested paper prototypes, click dummies, and preliminary app versions. Their engagement as well as our field observations allowed us to reveal and subsequently consider paramedic experience-based knowledge, and both were invaluable to designing and refining the app. We would not have gained this insight relying on literature and guidelines alone. For example, parents of sick toddlers must be asked whether they have a baby carrier to guarantee safe transport.

The engagement of paramedics with different professional experiences, ages, genders, and attitudes towards (digital) innovations was also valuable. In heterogeneous groups, we observed some controversial and inspiring discussions. For instance, paramedics in their early career expressed in the workshops that they adhered much more to guidelines (eg, they always followed the ABCDE scheme). In contrast, more experienced paramedics barely relied on these, which we also observed in role plays and during field observations. As our app needed to reflect the different ways paramedic users think and act, discussions helped to reconcile these different views.

The simulation of different rescue scenarios with both the paper prototypes and preliminary app versions helped us to check whether the features and content discussed in theory were of value and really needed in practice. The tests to explore how paramedics navigate through the app helped to build a user-friendly app.

The decisions regarding techniques used to explore the needs of paramedics were partly based on previous experience and on new questions and challenges that arose during the research and development process.

This open approach was especially helpful as it became clear that emergency medical services is an area that is also characterized by tacit knowledge and experience that is difficult to verbalize. Additionally, we discovered that the paramedics themselves often had different approaches when engaged in developing and testing activities. Some paramedics found it easier to reflect on personally experienced rescue missions with foreign-language patients and to assess their own behavior in these situations. For others, role play turned out to be a very effective tool as the scenarios helped paramedics imagine a concrete situation and this approach also involved more reserved participants. Paramedics became their own role model, and in some simulations, they acted contrary to their personal expectations. Through their involvement, they discovered processes or behavioral patterns they were previously unaware

of, for example gestures used automatically to enhance communication. Role play allowed both the paramedics and researchers to discover the paramedics' tacit knowledge of their work. In discussions subsequent to role play, paramedics would reflect on the communication process and combine their opinions in a very productive way.

The paramedics that tested the final app preferred to report their experience with it directly to our study nurse rather than using the messenger channel. These paramedics were familiar with the functionality and the content of the app and also provided peer teaching in the first phase of implementation.

## Discussion

### Principal Findings

In an action-oriented participatory approach, we successfully developed a digital communication tool to overcome language barriers in prehospital emergency care together with the end users — paramedics from 4 rescue stations. As the software designers took part in this process from the beginning, they gained background knowledge and were able to understand and discuss paramedics' needs and expectations regarding the functions requested. Pre-assumptions were challenged, which led to a greater understanding of the field. Communications and collaboration are challenging, especially if perspectives are contrary or if requirements cannot be met. Still, this complex approach ensured that the communication tool met the paramedics' needs, as it considered their perspectives informed by the problems they experienced in their day-to-day work. This is likely to increase the success of implementation in the longer term. In contrast to existing translation tools such as Google Translate, the developed app accommodates the needs of rescue service operations and enables quick and easy handling. It is independent of cell phone network coverage, and the translations were subject to an extensive quality assurance process. The possibility of documenting the communication process and patient responses are further aspects not offered by other solutions to date. Additionally, the app complies with current data protection and security standards.

Over the course of this study, it became apparent that a by-product of the app development was that it raised awareness among the participating paramedics of the communication barriers with foreign-language patients and the effect this has on the provision of care. It made paramedics think about their role and their previous behavior and attitudes, and they discussed these issues with their colleagues.

### Recommendations

This experience is captured in the following sections, and from it emerges a set of recommendations for future medical software development.

#### *End Users Are Experts of Their Work Environment*

Paramedics generated ideas concerning the app content and structure, user interface, and technical requirements. This ongoing user input in iterative cycles helped researchers and the software designers understand paramedics' needs. We experienced that end users were able to reflect on highly

complex issues. Thus, end users need to be taken seriously as experts of their work environment. In line with previous research [22,29], we argue that any new digital interventions in health care should actively involve end users in the design process, and we additionally recommend engaging software developers from the earliest stages of the project.

### ***A Participatory Approach Requires Time, Patience, and Humility***

During such an open process, unexpected issues might and can be expected to arise. It is essential to try not to anticipate the results, but to be open minded to any issues that come up. Planning a sufficiently long period for user testing of mock-ups facilitates input, feedback, and time to make specific adjustments prior to implementation.

### ***A Complex Intervention Benefits From a Range of Development Techniques***

The simulation of rescue scenarios with prototypes and other mock-ups was very helpful to structure and improve the content. Paramedic feedback on the user interface of the early app versions as well as the “think-aloud tasks” helped to build a user-friendly app. Therefore, we would encourage researchers to try a variety of approaches to foster development activities.

### ***The Development of a Digital Intervention for Health Care Requires a Multidisciplinary Team***

Working in a multidisciplinary team of clinical scientists, software developers, designers, and cultural scientists was crucial to meeting the challenge of building a communication tool suitable for paramedics in emergency medical services. We therefore recommend establishing a multidisciplinary team, ensuring that both clinical and technical competences are present.

### ***Sincerity is the Best Policy***

We were unable to fulfil all the needs identified during the development process. However, as we discussed, conflicting views on functionality with paramedics and explaining why we could not include certain functions did not diminish but rather enhanced their commitment. Should this be due to a lack of resources, we suggest having users prioritize their ideas and consider whether to incorporate disregarded functions in future versions. We also recommend discussing openly why some user-requested functions are not incorporated.

### ***Translation Poses the Risk of Altering the Intended Concepts and Meanings to Be Relayed***

Even though potential ambiguities were meant to be indicated in the final version of the phrases, some translations were found

to be unintentionally misleading. We recommend reviewing all phrases before and after translations with a person(s) proficient in the respective languages and medical professionals. Various rounds of forward translation or forward- and backward translation increase the validity of the translation and reduce the risk of mistranslation [30].

### ***Anticipate Future Ways to Use the New Technology***

The software developers contributed valuable ideas regarding further opportunities and challenges that may arise after app deployment. For example, they considered possibilities to directly transfer the documented data into electronic health records. The app was optimized so it could be run on different devices (eg, tablets) and different operating systems, to ease adoption and dissemination. In order to minimize potential compatibility issues, we recommend thinking about potential future use at an early stage and to attempt to accommodate this, where it is possible to do so, but not necessarily including it functionally.

### **Conclusions**

Clinician researchers, paramedics, software designers, and translators alone could not have built and created this comprehensive app. The expertise of this multiprofessional team was key to the success of developing this app. Digitization in today’s health care needs creative alliances between practitioners and developers: Bringing end users and software designers together, encouraging them to discuss ideas and opinions, giving a voice (and a say) to end users, and acknowledging their views on content and functionality are vital for developing digital innovations in health care.

### **Outlook**

Currently, the app is utilized in the work routines in the participating rescue stations. Within an interventional trial, we will assess the app’s usability, acceptance, effect on communication with foreign-language patients, and effect on information collection. We hope that our tool contributes to better and safer provision of paramedic care for foreign-language patients. We also plan to have the app evaluated and discussed by patient representatives with language skills in the supported languages.

Due to the COVID-19 pandemic, we upgraded the app with COVID-19-relevant questions covering associated symptoms and important information as well as phrases to communicate self-protection measures. The marketed version of the app has been released and is available for Android and iOS as “aidminutes.rescue (COVID-19)” in various app stores.

## **Acknowledgments**

We sincerely thank the paramedics of the rescue stations in Königslutter, Wendhausen, Helmstedt, and Braunschweig and the students and instructors of the School for Paramedics in Braunschweig for their participation. We are grateful to the team at aidminutes GmbH, in particular the app main software designer, Philipp Geisler, for their commitment. We acknowledge the support of the Malteser Hilfsdienst gGmbH, the district of Helmstedt, and the Fire Department of the City of Braunschweig.

The study was funded by the Federal Office for Agriculture and Food (BLE) of the German Federal Ministry of Food and Agriculture and Ministry for Social Affairs, Health and Equal Opportunities and by the subsidy guidelines for “Social innovation” of the European Social Fund (ESF). The app development by aidminutes GmbH was paid by these awards. The development of the marketed and publicly available version “aidminutes.rescue (COVID-19)” was accomplished beyond the actual project funding and without remuneration for aidminutes GmbH. The authors have no financial or other ties to aidminutes GmbH.

### Authors' Contributions

EN, JS, and FM were involved in the planning and realization of the study. EN and JS analyzed the data. EN was mainly responsible for drafting and writing the manuscript, JS and FM revised, and EN finalized, the manuscript. All authors have read and approved the final version of the manuscript.

### Conflicts of Interest

The authors declare no conflict of interest. The app development by aidminutes GmbH was paid by the awards listed above. The development of the marketed and publicly available version “aidminutes.rescue (COVID-19)” was accomplished beyond the actual project funding and without remuneration for aidminutes GmbH. The authors have no financial or other ties to aidminutes GmbH.

### Multimedia Appendix 1

Metamorphosis of app development and impressions of the participative development techniques. A-D Metamorphosis of app development with phrase construction and grouping (A), paper-based app (B), first click dummies (C) and final app (English interface) (D). E-H is showing participative approaches with workshops with paramedics in the rescue stations (E, F), and case simulations with paper-based app (G) and preliminary versions (H).

[[PNG File , 2989 KB - mhealth\\_v9i4e21586\\_app1.png](#) ]

### Multimedia Appendix 2

Links to the app stores (Android and iOS).

[[DOCX File , 59 KB - mhealth\\_v9i4e21586\\_app2.docx](#) ]

### References

1. Flores G. The impact of medical interpreter services on the quality of health care: a systematic review. *Med Care Res Rev* 2005 Jun;62(3):255-299. [doi: [10.1177/1077558705275416](#)] [Medline: [15894705](#)]
2. Karliner LS, Jacobs EA, Chen AH, Mutha S. Do professional interpreters improve clinical care for patients with limited English proficiency? A systematic review of the literature. *Health Serv Res* 2007 Apr;42(2):727-754 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00629.x](#)] [Medline: [17362215](#)]
3. Bard MR, Goettler CE, Schenarts PJ, Collins BA, Toschlog EA, Sagraves SG, et al. Language barrier leads to the unnecessary intubation of trauma patients. *Am Surg* 2004 Sep;70(9):783-786. [Medline: [15481294](#)]
4. Grow RW, Sztajnkrzyer MD, Moore BR. Language barriers as a reported cause of prehospital care delay in Minnesota. *Prehosp Emerg Care* 2008;12(1):76-79. [doi: [10.1080/10903120701709878](#)] [Medline: [18189182](#)]
5. Müller F, Hummers E, Noack EM. Medical Characteristics of Foreign Language Patients in Paramedic Care. *Int J Environ Res Public Health* 2020 Aug 30;17(17) [FREE Full text] [doi: [10.3390/ijerph17176306](#)] [Medline: [32872563](#)]
6. Alm-Pfrunder AB, Falk A, Vicente V, Lindström V. Prehospital emergency care nurses' strategies while caring for patients with limited Swedish-English proficiency. *J Clin Nurs* 2018 Oct;27(19-20):3699-3705. [doi: [10.1111/jocn.14484](#)] [Medline: [29679408](#)]
7. Tate RC, Hodkinson PW, Meehan-Coussee K, Cooperstein N. Strategies Used by Prehospital Providers to Overcome Language Barriers. *Prehosp Emerg Care* 2016;20(3):404-414. [doi: [10.3109/10903127.2015.1102994](#)] [Medline: [26807873](#)]
8. Diamond LC, Schenker Y, Curry L, Bradley EH, Fernandez A. Getting by: underuse of interpreters by resident physicians. *J Gen Intern Med* 2009 Feb;24(2):256-262 [FREE Full text] [doi: [10.1007/s11606-008-0875-7](#)] [Medline: [19089503](#)]
9. Hsieh E. Not just "getting by": factors influencing providers' choice of interpreters. *J Gen Intern Med* 2015 Jan;30(1):75-82 [FREE Full text] [doi: [10.1007/s11606-014-3066-8](#)] [Medline: [25338731](#)]
10. Würth KM, Reiter-Theil S, Langewitz W, Schuster S. “Getting by” in a Swiss Tertiary Hospital: the Inconspicuous Complexity of Decision-making Around Patients’ Limited Language Proficiency. *J GEN INTERN MED* 2018 Aug 24;33(11):1885-1891 [FREE Full text] [doi: [10.1007/s11606-018-4618-0](#)] [Medline: [30143979](#)]
11. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K. Development of machine translation technology for assisting health communication: A systematic review. *J Biomed Inform* 2018 Sep;85:56-67 [FREE Full text] [doi: [10.1016/j.jbi.2018.07.018](#)] [Medline: [30031857](#)]
12. Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuracy. *BMJ* 2014 Dec 15;349:g7392 [FREE Full text] [doi: [10.1136/bmj.g7392](#)] [Medline: [25512386](#)]

13. Müller F, Chandra S, Furaijat G, Kruse S, Waligorski A, Simmenroth A, et al. A Digital Communication Assistance Tool (DCAT) to Obtain Medical History from Foreign-Language Patients: Development and Pilot Testing in a Primary Health Care Center for Refugees. *Int J Environ Res Public Health* 2020 Feb 20;17(4) [[FREE Full text](#)] [doi: [10.3390/ijerph17041368](https://doi.org/10.3390/ijerph17041368)] [Medline: [32093286](https://pubmed.ncbi.nlm.nih.gov/32093286/)]
14. Furaijat G, Kleinert E, Simmenroth A, Müller F. Implementing a digital communication assistance tool to collect the medical history of refugee patients: DICTUM Friedland - an action-oriented mixed methods study protocol. *BMC Health Serv Res* 2019 Feb 06;19(1):103 [[FREE Full text](#)] [doi: [10.1186/s12913-019-3928-1](https://doi.org/10.1186/s12913-019-3928-1)] [Medline: [30728030](https://pubmed.ncbi.nlm.nih.gov/30728030/)]
15. Spechbach H, Gerlach J, Mazouri Karker S, Tsourakis N, Combesure C, Bouillon P. A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study. *JMIR Med Inform* 2019 May 07;7(2):e13167 [[FREE Full text](#)] [doi: [10.2196/13167](https://doi.org/10.2196/13167)] [Medline: [31066702](https://pubmed.ncbi.nlm.nih.gov/31066702/)]
16. Wherton J, Sugarhood P, Procter R, Hinder S, Greenhalgh T. Co-production in practice: how people with assisted living needs can help design and evolve technologies and services. *Implement Sci* 2015 May 26;10:75 [[FREE Full text](#)] [doi: [10.1186/s13012-015-0271-8](https://doi.org/10.1186/s13012-015-0271-8)] [Medline: [26004047](https://pubmed.ncbi.nlm.nih.gov/26004047/)]
17. Robert G. Participatory action research: using experience-based co-design to improve the quality of healthcare services. In: Ziebland S, Coulter A, Calabrese JD, Lcock L, editors. *Understanding and using health experiences: Improving patient care*, 1st edition. Oxford, England: Oxford University Press; 2013:138-148.
18. van Limburg M, Wentzel J, Sanderman R, van Gemert-Pijnen L. Business Modeling to Implement an eHealth Portal for Infection Control: A Reflection on Co-Creation With Stakeholders. *JMIR Res Protoc* 2015 Aug 13;4(3):e104 [[FREE Full text](#)] [doi: [10.2196/resprot.4519](https://doi.org/10.2196/resprot.4519)] [Medline: [26272510](https://pubmed.ncbi.nlm.nih.gov/26272510/)]
19. Noack EM, Kleinert E, Müller F. Overcoming language barriers in paramedic care: a study protocol of the interventional trial 'DICTUM rescue' evaluating an app designed to improve communication between paramedics and foreign-language patients. *BMC Health Serv Res* 2020 Mar 18;20(1):223 [[FREE Full text](#)] [doi: [10.1186/s12913-020-05098-5](https://doi.org/10.1186/s12913-020-05098-5)] [Medline: [32183775](https://pubmed.ncbi.nlm.nih.gov/32183775/)]
20. Zojer E, Faul E, Mayer H. Aktionsforschung — „Be part of it“. *ProCare* 2013 Nov 14;18(9):12-16. [doi: [10.1007/s00735-013-0185-0](https://doi.org/10.1007/s00735-013-0185-0)]
21. Lewin K. Action research and minority problems. *Journal of Social Issues* 1946;2(4):34-46. [doi: [10.1111/j.1540-4560.1946.tb02295.x](https://doi.org/10.1111/j.1540-4560.1946.tb02295.x)]
22. Clemensen J, Larsen SB, Kyng M, Kirkevold M. Participatory design in health sciences: Using cooperative experimental methods in developing health services and computer technology. *Qual Health Res* 2007 Jan;17(1):122-130. [doi: [10.1177/1049732306293664](https://doi.org/10.1177/1049732306293664)] [Medline: [17170250](https://pubmed.ncbi.nlm.nih.gov/17170250/)]
23. Warren LR, Harrison M, Arora S, Darzi A. Working with patients and the public to design an electronic health record interface: a qualitative mixed-methods study. *BMC Med Inform Decis Mak* 2019 Dec 03;19(1):250 [[FREE Full text](#)] [doi: [10.1186/s12911-019-0993-7](https://doi.org/10.1186/s12911-019-0993-7)] [Medline: [31795998](https://pubmed.ncbi.nlm.nih.gov/31795998/)]
24. Bustamante Duarte AM, Brendel N, Degbelo A, Kray C. Participatory Design and Participatory Research. *ACM Trans. Comput.-Hum. Interact* 2018 Feb 24;25(1):1-39. [doi: [10.1145/3145472](https://doi.org/10.1145/3145472)]
25. Victor CR, Peacock JL, Chazot C, Walsh S, Holmes D. Who calls 999 and why? A survey of the emergency workload of the London Ambulance Service. *J Accid Emerg Med* 1999 May;16(3):174-178 [[FREE Full text](#)] [doi: [10.1136/emj.16.3.174](https://doi.org/10.1136/emj.16.3.174)] [Medline: [10353041](https://pubmed.ncbi.nlm.nih.gov/10353041/)]
26. Møller TP, Ersbøll AK, Tolstrup JS, Østergaard D, Viereck S, Overton J, et al. Why and when citizens call for emergency help: an observational study of 211,193 medical emergency calls. *Scand J Trauma Resusc Emerg Med* 2015 Nov 04;23:88 [[FREE Full text](#)] [doi: [10.1186/s13049-015-0169-0](https://doi.org/10.1186/s13049-015-0169-0)] [Medline: [26530307](https://pubmed.ncbi.nlm.nih.gov/26530307/)]
27. Soar J, Nolan JP, Böttiger BW, Perkins GD, Lott C, Carli P, Adult advanced life support section Collaborators. European Resuscitation Council Guidelines for Resuscitation 2015: Section 3. Adult advanced life support. *Resuscitation* 2015 Oct;95:100-147. [doi: [10.1016/j.resuscitation.2015.07.016](https://doi.org/10.1016/j.resuscitation.2015.07.016)] [Medline: [26477701](https://pubmed.ncbi.nlm.nih.gov/26477701/)]
28. Marx J, Hockberger R, Walls R, Adams J, Rosen P, editors. *Rosen's emergency medicine: Concepts and clinical practice*, 7th edition. Philadelphia, PA: Mosby/Elsevier; 2010.
29. Castensøe-Seidenfaden P, Reventlov Husted G, Teilmann G, Hommel E, Olsen BS, Kensing F. Designing a Self-Management App for Young People With Type 1 Diabetes: Methodological Challenges, Experiences, and Recommendations. *JMIR Mhealth Uhealth* 2017 Oct 23;5(10):e124 [[FREE Full text](#)] [doi: [10.2196/mhealth.8137](https://doi.org/10.2196/mhealth.8137)] [Medline: [29061552](https://pubmed.ncbi.nlm.nih.gov/29061552/)]
30. Brammen D, Greiner F, Dormann H, Mach C, Wrede C, Ballaschk A, et al. Lessons learned in applying the International Society for Pharmacoeconomics and Outcomes Research methodology to translating Canadian Emergency Department Information System Presenting Complaints List into German. *Eur J Emerg Med* 2018 Aug;25(4):295-299 [[FREE Full text](#)] [doi: [10.1097/MEJ.0000000000000450](https://doi.org/10.1097/MEJ.0000000000000450)] [Medline: [28145941](https://pubmed.ncbi.nlm.nih.gov/28145941/)]

*Edited by G Eysenbach; submitted 20.07.20; peer-reviewed by T Matsuyama, E Ray Chaudhuri, C Berdahl; comments to author 05.10.20; revised version received 29.11.20; accepted 06.02.21; published 14.04.21.*

*Please cite as:*

Noack EM, Schulze J, Müller F

*Designing an App to Overcome Language Barriers in the Delivery of Emergency Medical Services: Participatory Development Process*  
*JMIR Mhealth Uhealth* 2021;9(4):e21586

URL: <https://mhealth.jmir.org/2021/4/e21586>

doi: [10.2196/21586](https://doi.org/10.2196/21586)

PMID: [33851933](https://pubmed.ncbi.nlm.nih.gov/33851933/)

©Eva Maria Noack, Jennifer Schulze, Frank Müller. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 14.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Baby Buddy App for Breastfeeding and Behavior Change: Retrospective Study of the App Using the Behavior Change Wheel

Loretta M Musgrave<sup>1,2</sup>, MEd; Alison Baum<sup>3</sup>, MSc; Nilushka Perera<sup>3</sup>, MPH; Caroline SE Homer<sup>1,4</sup>, PhD; Adrienne Gordon<sup>2,5</sup>, PhD

<sup>1</sup>Centre for Midwifery, Child and Family Health, Faculty of Health, University of Technology Sydney, Ultimo NSW, Australia

<sup>2</sup>Charles Perkins Centre, Faculty of Medicine and Health, University of Sydney, Camperdown NSW, Australia

<sup>3</sup>Best Beginnings, London, United Kingdom

<sup>4</sup>Burnet Institute, Melbourne VIC, Australia

<sup>5</sup>Royal Prince Alfred Hospital, Sydney Local Health District, NSW Health, Camperdown NSW, Australia

**Corresponding Author:**

Loretta M Musgrave, MEd

Centre for Midwifery, Child and Family Health

Faculty of Health

University of Technology Sydney

Building 10, Level 11

235 Jones St

Ultimo NSW, 2006

Australia

Phone: 61 (02) 9514 5069

Email: [loretta.musgrave@uts.edu.au](mailto:loretta.musgrave@uts.edu.au)

## Abstract

**Background:** Breastfeeding plays a major role in the health of mothers and babies and has the potential to positively shape an individual's life both in the short and long term. In the United Kingdom (UK), although 81% of women initiate breastfeeding, only 1% of women breastfeed exclusively to 6 months as recommended by the World Health Organization. In the UK, women who are socially disadvantaged and younger are less likely to breastfeed at 6 to 8 weeks postpartum. One strategy that aims to improve these statistics is the Baby Buddy app, which has been designed and implemented by the UK charity Best Beginnings to be a universal intervention to help reduce health inequalities, including those in breastfeeding.

**Objective:** This study aimed to retrospectively examine the development of Baby Buddy by applying the Behavior Change Wheel (BCW) framework to understand how it might increase breastfeeding self-efficacy, knowledge, and confidence.

**Methods:** Retrospective application of the BCW was completed after the app was developed and embedded into maternity services. A three-stage process evaluation used triangulation methods and formalized tools to gain an understanding of the potential mechanisms and behaviors used in apps that are needed to improve breastfeeding rates in the UK. First, we generated a behavioral analysis by mapping breastfeeding barriers and enablers onto the Capability, Opportunity, and Motivation-Behavior (COM-B) system using documents provided by Best Beginnings. Second, we identified the intervention functions and policy categories used. Third, we linked these with the behavior change techniques identified in the app breastfeeding content using the Behavior Change Techniques Taxonomy (BCTTv1).

**Results:** Baby Buddy is a well-designed platform that could be used to change breastfeeding behaviors. Findings from stage one showed that Best Beginnings had defined breastfeeding as a key behavior requiring support and demonstrated a thorough understanding of the context in which breastfeeding occurs, the barriers and enablers of breastfeeding, and the target actions needed to support breastfeeding. In stage two, Best Beginnings had used intervention and policy functions to address the barriers and enablers of breastfeeding. In stage three, Baby Buddy had been assessed for acceptability, practicability, effectiveness, affordability, safety, and equity. Several behavior change techniques that could assist women with decision making around breastfeeding (eg, information about health consequences and credible sources) and possibly affect attitudes and self-efficacy were identified. Of the 39 videos in the app, 19 (49%) addressed physical capabilities related to breastfeeding and demonstrated positive breastfeeding behaviors.



**Conclusions:** Applying a theoretical framework retrospectively to a mobile app is possible and results in useful information to understand potential health benefits and to inform future development. Future research should assess which components and behavioral techniques in the app are most effective in changing behavior and supporting breastfeeding.

(*JMIR Mhealth Uhealth* 2021;9(4):e25668) doi:[10.2196/25668](https://doi.org/10.2196/25668)

## KEYWORDS

breastfeeding; app; digital health; smartphone app; behavior change wheel; digital behavior change intervention

## Introduction

A healthy start to life is crucial for improving life-long health outcomes [1,2]. Despite universal public funding for pregnancy care and targeted antenatal and postnatal programs, the United Kingdom (UK) has large inequalities in perinatal outcomes for women and children from minority ethnic communities, those who are socially disadvantaged, or those who become pregnant in their teenage years [3]. Breastfeeding is well recognized to positively impact and shape the lives of both the mother and baby in the short and long term. Global scaling up of breastfeeding interventions is needed to improve the rates of breastfeeding in all countries, which includes the provision of support to all women [1,2].

Breastmilk is nutritionally balanced and helps protect infants and children from infections [1]. There are risks associated with not breastfeeding in high-income, middle-income, and low-income countries [1]. A meta-analysis of six high-quality studies showed that “ever breastfeeding” (infants who have breastfed at least once) was associated with a 36% reduction in sudden infant death (95% CI 19%-49%) [2]. Breastfed babies have a lower chance of childhood leukemia and allergies, and are less likely to develop diabetes or become overweight when they are older [4]. Breastfeeding also benefits mothers, and it is associated with a lower risk of developing breast and ovarian cancer, osteoporosis, diabetes, and cardiovascular disease [2]. A longer period of breastfeeding is also associated with a reduction in the mother's odds of overweight or obesity (95% CI 22-30) [4].

The UK National Infant Feeding Survey (2010) showed that although 81% of women initiated breastfeeding, 34% of babies received any breastmilk at 6 months (only 1% were exclusively breastfed), and the country ranks lowest in the world for breastfeeding at 12 months of age [2,3]. The most recent aggregate breastfeeding rate for England (Quarter 3 of 2019/20) at 6 to 8 weeks was 48.2% (CI 47.9%-48.5%) [5]. As a response to low breastfeeding rates, the UK Public Health England in collaboration with UNICEF UK, has produced several policies and resources in line with the “baby friendly initiative.” It is hoped that initiatives that promote breastfeeding will augment women's and children's health and support maternal-infant bonding [6,7].

In 2007, Best Beginnings charity in the UK co-designed digital video discs (DVDs) to support breastfeeding initiation, motivation, and duration, with a focus on benefits and acknowledgement of challenges. The resources were developed with parents, the UK Department of Health, and UNICEF UK. Since the 2008 launch, over 2 million copies of the DVD have

been distributed. In 2014, with changing technology, the charity embedded this breastfeeding content into Baby Buddy, a smartphone app. Pregnant women are now more likely to find pregnancy apps useful sources of information and support compared with DVDs or written material [8-12]. This trend toward the use of smartphones provides an opportunity to reach those women who are less likely to engage with health care providers or are yet to do so [13,14].

Baby Buddy was designed to focus on the window of opportunity from preconception to 6 months of age, in which the foundations for a healthy childhood are laid [15]. The app is free, available on the National Health Service Library, embedded into maternity and early care pathways, and endorsed by organizations, including the Royal College of Midwives and the Royal College of Obstetricians and Gynaecologists, and it can be easily accessed on both Android and iOS devices. Baby Buddy is intended to be used by parents of all backgrounds and to be particularly engaging for those who may have difficulty connecting with health services owing to language, age, culture, or socioeconomic barriers. Baby Buddy has been designed to appeal to younger women and includes a user-designed interactive avatar as a “gaming” element. The app aims to build confidence and self-efficacy and promote good parental-infant bonding and attachment. It contains over 300 videos, including all videos from the “From bump to breastfeeding” DVD, and provides engaging and interactive daily information to support healthy behaviors including breastfeeding. The app intends to enhance the link between parents and health care providers and promotes better engagement, communication, and shared decision making with parents [16].

The most recent published evaluation of Baby Buddy, the BaBBLes study (Bumps and Babies Longitudinal Study), measured maternal self-efficacy as the primary outcome. The authors found that there were no differences in maternal self-efficacy outcomes. However, they did perform a post-hoc analysis of breastfeeding and documented a significant increase in “any breastfeeding” at 1 month (odds ratio [OR] 3.08, 95% CI 1.49-6.35) and in “exclusive breastfeeding” at 3 months (OR 1.79, 95% CI 1.02-3.16) [16]. Further data from Norfolk did demonstrate an increase in maternal self-efficacy for parents using the Baby Buddy app [17].

With this data demonstrating potential behavior change and increased breastfeeding with the use of the Baby Buddy app, further understanding was sought regarding which components of the design and development of the app might have contributed to these results. The Behavior Change Wheel (BCW) and the associated Behavior Change Techniques Taxonomy (BCTTv1) provide a systematic approach that acknowledges the importance of behavioral theory in the design and evaluation of

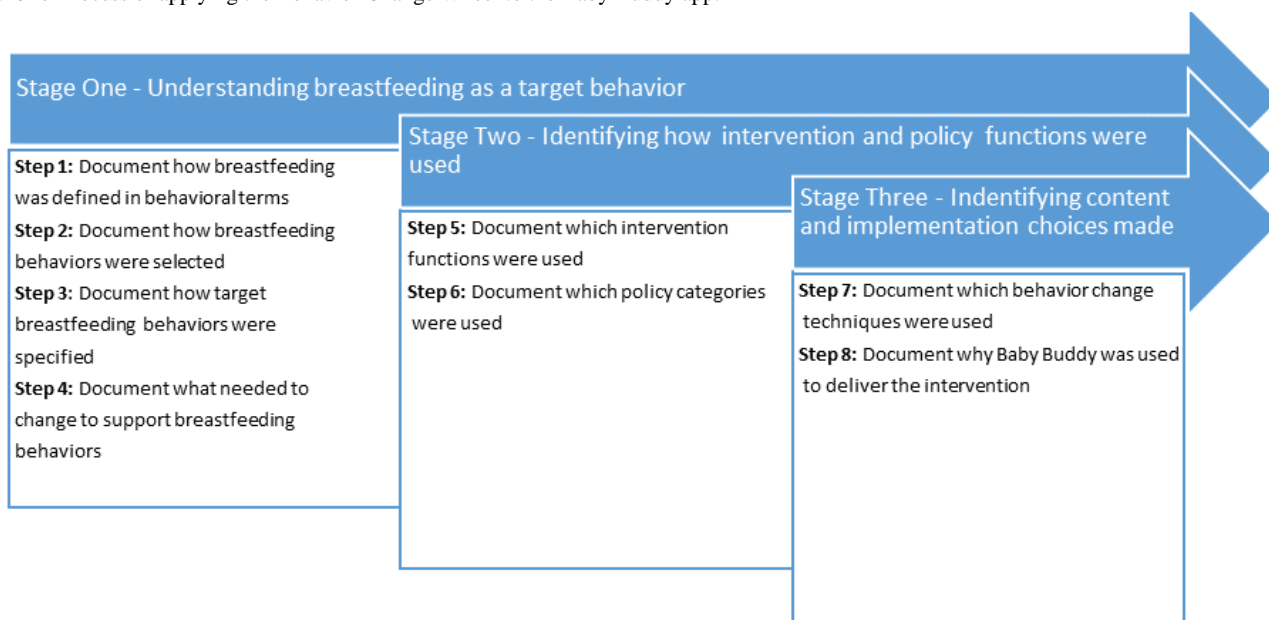
interventions. The BCW has three interrelated concentric layers. The inner layer (Capability, Opportunity, and Motivation-Behavior [COM-B] model) helps understand the behavior that needs to be changed. The middle layer consists of the following possible interventions that could be used to facilitate behavior change: restrictions, education, persuasion, incentivization, coercion, training, enablement, modeling, and environmental restructuring. The outer layer of the wheel assists in identifying which policy opportunities could be utilized to support the delivery of the chosen interventions [18]. Finally, the BCTTv1 is a complementary tool that helps further identify which behavior change techniques could help deliver the intervention functions identified [18]. The BCW has previously been retrospectively applied to other mobile health interventions successfully [19-21]. This study aimed to retrospectively examine the development of Baby Buddy and apply the BCW framework to understand how it might increase breastfeeding self-efficacy, knowledge, and confidence.

## Methods

### Overview

We evaluated the development of Baby Buddy with the BCW and its associated taxonomy using a prespecified three-stage process (Figure 1). The research was conducted between November 2017 and December 2018. The research team was given access to all reports, market research, and interview and focus group findings prepared by Best Beginnings to inform the design of Baby Buddy (Multimedia Appendix 1). Guide books containing worksheets were used to deconstruct and retrospectively analyze the development process and the breastfeeding components within the Baby Buddy app [18,22]. Data extraction was performed by one reviewer (LMM) and then checked by a second (AG). They met fortnightly to share and discuss the findings. This was achieved by cross-checking coding, interpretation, and mapping. Any discrepancy was resolved by discussion, and further analysis or content review was undertaken if necessary.

Figure 1. Process of applying the Behavior Change Wheel to the Baby Buddy app.



### Stage One: Understanding Breastfeeding as a Target Behavior

This stage aimed to assess the in-depth understanding of breastfeeding as a target behavior in the development of the app and the context in which it occurs. Barriers and enablers to target behaviors were identified in the provided data (survey, interview, and focus group reports), and then, these were mapped to the COM-B tool [18].

### Stage Two: Identifying How Intervention and Policy Functions Were Used

This stage determined the aspects included in Baby Buddy and if they could influence breastfeeding behavior. The middle layer of the BCW was used by the research team to map which “intervention” components could address the barriers and enablers to breastfeeding (restrictions, education, persuasion, incentivization, coercion, training, enablement, modeling, and

environmental restructuring) [18]. We then coded these findings using the Theoretical Domains Framework (TDF) (knowledge, cognitive and interpersonal skills, memory, attention and decision-making processes, optimism, beliefs about consequences, intentions, goals, emotions, and social influences) [18]. The outer layer of the BCW was used to map policy categories (eg, policies, guidelines, fiscal measures, service provision, legislation, regulation, communication, and environmental opportunities) [18].

### Stage Three: Identifying Content and Implementation Choices Made

This stage identified evidence of the use of behavioral change techniques (BCTs) within the design of Baby Buddy. We used the APEASE criteria as defined in the BCW (affordability, practicability, effectiveness, affordability, safety, and equity) [18,22]. These steps provided insights into how the content was developed and implemented and to understand the choices made

by Best Beginnings as the project progressed. We also described the “active” ingredients that were used in the breastfeeding intervention using the BCTTv1 tool [18,22]. To do this, we viewed and reviewed 39 videos, eight glossary words (“What does that mean?”), and 20 Baby Buddy-generated responses to breastfeeding questions (“Ask me”). We marked the BCTTv1 tool for each technique found in each piece of information reviewed (videos, glossary words, and generated responses).

## Results

### Stage One: Understanding Breastfeeding as a Target Behavior

#### Step 1: How Breastfeeding was Defined in Behavioral Terms

Best Beginnings defined breastfeeding as a key behavior requiring more support and demonstrated a thorough understanding of the context in which breastfeeding occurs based on the following sources of evidence: (1) The Infant Feeding Survey (2010) [3]; (2) World Health Organization Global Strategy for Infant and Young Child Feeding (Breastfeeding Manifesto) [23]; (3) Tackling health inequalities in infant and maternal health outcomes [15]; (4) Focus On: A Proportionate Approach to Priority Populations [24]; (5) Fair Society, Healthy Lives – Strategic Review of Health Inequalities in England post-2010 [25]; and (6) The Foundation Years: preventing poor children becoming poor adults [26].

**Table 1.** Specifying breastfeeding as a target behavior [22].

Question	Response
Who needs to perform the behavior?	Women, with a focus on young women under 25 years intending to breastfeed
What does the person need to do differently to achieve the desired change?	Offer breast first
When will they do it?	Within the first hour of birth and then for every feed demanded
Where will they do it?	At the birthplace and then anywhere they choose to feed the infant
How often will they do it?	Every feed
With whom will they do it?	With the support of staff initially and then independently with the support of family and friends or professionals if required

#### Step 4: Changes Needed to Support Breastfeeding Behaviors

We found evidence to support that the constructs of capability, opportunity, and motivation were explored as described below.

##### Physical and Psychological Capability

Best Beginnings explored social norms, peer influence, and the value of social support in sustaining breastfeeding. For example, women were asked to discuss breastfeeding in the context of their roles in their families, the presence or absence of support, the influences of cultural values, and the impact of migration, isolation, and loneliness. Peer and clinical support, demonstrations, practice, and feedback were seen as important to enable women to breastfeed. Perceived barriers, such as difficulties positioning and attaching, low milk production (physical capability), fear of failure, and anxiety/depression

#### Step 2: How Breastfeeding Behaviors Were Selected

Best Beginnings selected specified target actions that were needed to support breastfeeding. They undertook extensive consultation with stakeholders, including the UK Department of Health, UNICEF, and women and their families. A multidisciplinary team approach was adopted in the creation of the steering committee. The following six target behaviors to support breastfeeding were identified as a priority by Best Beginnings: (1) Advising on commencing breastfeeding; (2) Giving information on correct positioning and attachment for breastfeeding; (3) Knowing how to express breast milk; (4) Knowing what is normal in the first few months of breastfeeding; (5) Knowing how to overcome breastfeeding challenges; and (6) Planning to breastfeed for 6 months or more.

#### Step 3: How Target Breastfeeding Behaviors Were Specified

Breastfeeding behaviors were described with who, what, when, where, how often, and with whom (Table 1). Best Beginnings utilized mixed method techniques to better understand the barriers and enablers affecting inequity, disparity, and intergenerational disadvantage (Multimedia Appendix 1). Health care professionals, parents, and families were engaged as co-creators at all stages and were instrumental in app development, implementation, evaluation, and promotion [16,27-30].

(psychological capability), were identified as needing to be addressed by the intervention functions (Multimedia Appendix 2).

##### Physical and Social Opportunity

Support was identified as the primary enabler for both physical and social opportunity to breastfeed. Clinical/specialist, peer, community, and technology supports (apps, social media, and online resources) were documented as facilitators for breastfeeding. Best Beginnings sought to understand environmental factors that may help, interfere, or prevent breastfeeding efforts. Economic barriers and the physical environment were discussed, and there were several themes related to challenges in finding a way to initiate and maintain breastfeeding behaviors in the context of roles as employees, mothers, and partners (Multimedia Appendix 2).

### Reflective and Automatic Motivation

Reviewed data demonstrated that motivation is best facilitated by early planning, goal setting, and positive belief reinforcement. Peer support normalizes the challenges of breastfeeding and encourages self-determination. Best Beginnings documented support as crucial to helping alleviate negative thoughts or low confidence. Self-efficacy to change beliefs and habits, and low health literacy barriers were explored to assess the ability of individuals to act on health advice and planned care and to uncover culturally specific values that may improve interventions in specific target groups ([Multimedia Appendix 2](#)).

### Stage Two: Identifying How Intervention Functions Were Used

#### *Step 5: Intervention Functions That Were Used*

Intervention functions were able to be identified in the video content for Baby Buddy, which included the lead information and education resource within the app. The most common functions were education, training, and modeling. Mapping of the breastfeeding video content to the BCW (COM-B, TDF, and intervention functions) is shown in [Multimedia Appendix 2](#). Further analysis of each video containing breastfeeding content (39 videos) is shown in [Table 2](#). The complete analysis of all breastfeeding items, including eight glossary words (“What does that mean?”) and 20 Baby Buddy-generated responses to breastfeeding questions (‘Ask me’), using COM-B, is shown in [Multimedia Appendix 3](#).

**Table 2.** Mapping of breastfeeding video content to the Capability, Opportunity, and Motivation-Behavior (COM-B) tool.

Video title	Capability		Opportunity		Motivation	
	Physical	Psychological	Social	Physical	Reflective	Automatic
Breastfeeding as a young mum	Yes	Yes	Yes	Yes	Yes	Yes
A practical choice	Yes	Yes	Yes	Yes	Yes	Yes
Feelings about breastfeeding	Yes	Yes	Yes	Yes	Yes	Yes
What's so good about breastfeeding?	No	Yes	Yes	Yes	Yes	Yes
What if I bottle fed before?	Yes	Yes	Yes	Yes	Yes	Yes
Asking for help to get started	Yes	Yes	Yes	Yes	No	No
What will my partner think?	No	No	Yes	No	Yes	No
Your first milk - colostrum	Yes	Yes	No	No	Yes	Yes
Your baby's first feed	Yes	Yes	Yes	Yes	No	Yes
Skin to skin	Yes	Yes	Yes	Yes	No	Yes
Good positioning tips from a midwife	Yes	No	Yes	No	No	Yes
Getting the position right	Yes	No	No	Yes	No	Yes
Good positioning demonstration	Yes	Yes	No	Yes	Yes	No
Keeping your baby close	Yes	Yes	Yes	Yes	Yes	Yes
How dads can help? - Lenny	No	No	Yes	Yes	No	No
Breastfeeding out and about	No	Yes	Yes	No	Yes	Yes
When and how often should I feed my baby?	Yes	Yes	Yes	Yes	Yes	Yes
How dads can help? - Andy	No	No	Yes	Yes	Yes	No
Where can I find support?	Yes	Yes	Yes	Yes	Yes	No
Overcoming mastitis	Yes	Yes	Yes	Yes	Yes	Yes
Support from health professionals	Yes	Yes	Yes	Yes	Yes	Yes
Some common challenges	Yes	Yes	No	Yes	Yes	No
Good and bad attachment graphic	Yes	No	No	Yes	No	No
Breastfeeding to a year and beyond	Yes	Yes	Yes	Yes	Yes	Yes
Why breastfeed for at least six months?	Yes	Yes	Yes	Yes	Yes	Yes
Breastfeeding and weening	Yes	Yes	Yes	Yes	Yes	Yes
Why express?	Yes	Yes	Yes	Yes	Yes	Yes
How to hand express?	Yes	Yes	No	Yes	Yes	Yes
How to use a breast pump?	Yes	No	No	Yes	Yes	No
Expressing when you're back at work	Yes	Yes	No	Yes	Yes	No
Storing and using expressed breast milk	No	No	No	Yes	Yes	No
Early challenges with expressing milk	Yes	Yes	No	Yes	Yes	No
Your breast milk	Yes	Yes	Yes	Yes	Yes	Yes
How skin-to-skin contact can help you express?	Yes	Yes	Yes	Yes	Yes	Yes
Using a breast pump	Yes	No	Yes	Yes	Yes	No
Expressing with a breast pump and storing your milk	Yes	Yes	Yes	Yes	Yes	Yes
Colostrum - your baby's first food	Yes	Yes	No	Yes	No	Yes
Signs your baby is ready to feed independently	No	Yes	Yes	Yes	Yes	No
Breastfeeding twins or triplets	Yes	No	No	Yes	Yes	No

### ***Step 6: Policy Categories That Were Used***

Findings support that the Baby Buddy app and its embedding process were designed to complement maternity and postnatal health service and policy [28]. It has been endorsed by the Department of Health, Faculty of Public Health, Royal Colleges of Paediatrics and Child Health, obstetricians and gynecologists, midwives, psychiatrists, speech and language therapists, community practitioners, Health Visitors Association, and Institute of Health Visiting. The content of Baby Buddy was co-created with parents and in consultation with policy stakeholders, for example, representatives from Royal Colleges and the Department of Health. No content is uploaded to Baby Buddy until representatives of all partners have given their approval.

### **Stage Three: Identifying How Content and Implementation Choices Were Made**

#### ***Step 7: How Behavior Change Techniques Were Used***

Identification of BCTs was achieved by applying the BCTTv1 to the content of the app. After each piece of content was categorized using broad intervention categories, further analysis was carried out to identify exactly which BCTs were used (Multimedia Appendix 2). These were then documented and specific details were given. For example, of the 39 videos in the app, 19 (49%) addressed physical capabilities related to breastfeeding and demonstrated positive breastfeeding behaviors.

#### ***Step 8: Rationale For Using the Baby Buddy App as the Mode of Delivery***

The APEASE criteria were used to evaluate if Best Beginnings had undertaken activities to ascertain acceptability, practicability, effectiveness, affordability, safety, and equity when moving breastfeeding content to a mobile app. The evidence was analyzed and judged against the previous DVD-based breastfeeding intervention, “Bump to breastfeeding.” Baby Buddy met the APEASE criteria for a viable digital intervention suitable for further testing, development, and implementation (Multimedia Appendix 4). In addition, it was noted that in transitioning from DVD to a mobile app, Best Beginnings used the Kotter eight-step process to guide implementation. Kotter methodology, developed for change management, involves the following eight steps: (1) creating a sense of urgency, (2) building a guiding coalition, (3) forming strategic vision and initiatives, (4) enlisting a volunteer army, (5) enabling action by removing barriers, (6) generating short-term wins, (7) sustaining acceleration, and (8) instituting change [31,32].

## ***Discussion***

### **Principal Findings**

Baby Buddy maps retrospectively well to the BCW. This may explain why there have been positive results in recent studies [28,29]. Key factors that set the development of this particular pregnancy app apart from many others are the genuine co-design and the use of BCTs most obviously through the included video content.

The use of participatory engagement and co-creation methods in the development of Baby Buddy are two design techniques that have positively influenced decision making, attitudes, and self-efficacy concerning breastfeeding, particularly among those who are socially disadvantaged and younger. We identified several BCTs used in Baby Buddy that could assist women with decision making around breastfeeding (eg, BCT 5.1 Information about health consequences and 9.1 Credible source). BCTs that influence attitudes and self-efficacy were also identified (eg, BCT 5.3 Information about social and environmental consequences and 13.2 Framing/reframing).

### **Strengths and Limitations**

This study has several strengths. First, it was performed independent of the development team, using a best practice behavior change framework (BCW) as a guide. Second, content mapping to the BCW was conducted by two independent content experts (a midwife and a neonatologist). These two research team members located in Australia were not employed by Best Beginnings and did not have any financial incentive. Third, retrospective alignment of the BCW tools and BCTs enabled the research team to identify potential opportunities to use BCTs for the future development of Baby Buddy to increase effectiveness. Fourth, our study supports the work of Thomson and Crossland who conducted a mixed methods evaluation using the BCW to identify components that support infant feeding in North West UK [29]. They identified peer support as a facilitator for increasing mothers’ knowledge and building confidence [29]. Finally, we also identified the use of peer-to-peer content as beneficial for breastfeeding as it normalizes breastfeeding and encourages self-determination. Baby Buddy has both of these attributes in the content. Like the work of Crossland et al, our study concluded that Baby Buddy is a supportive parenting resource that could be scaled for impact [28].

A key limitation of this work is the retrospective application of the BCW. Retrospective mapping of the BCW to the app development process was complex and subjective, and relied on Best Beginnings providing multiple development documents. There was a large volume of qualitative reports supplied to us from Best Beginnings that had been collected from many sources and not presented with later academic review in mind.

Using the BCW has inherent coding, interpretation, and application limitations. However, like other studies, we do believe that there is benefit in “retrofitting” interventions to the BCW even though it may have not been used in the design phase [33-35]. Prospective analysis of the app development using the BCW and scientific research would potentially result in a higher quality behavior change intervention tool; however, Baby Buddy was not primarily designed to change behavior and was rather designed as a resource to inform and empower pregnant women.

A secondary limitation is that the evaluation tools we used were designed for text rather than video content. From our assessment, videos within an app appear to be a powerful influence to support behavior change in breastfeeding. The videos take a “show how” approach rather than a didactic “tell to” approach and feature a mixture of experts, support parents, and peer-to-peer voices. However, as the BCW tools were not designed for video discourse analysis specifically, they may

miss some of the nuances in video content (eg, gesturing, body language, and tone). Our findings have identified potential areas for improvement in future iterations of the app, and this is useful information given that the app is constantly being improved.

### Conclusion

Our work highlights that applying a theoretical framework retrospectively to a mobile health app is possible and results in useful information to understand potential health benefits and

to inform future development. To assess the true impact of behavior change frameworks in the design of mobile health apps, high-quality research that measures formative, process, and clinical outcomes for health behaviors is needed. Further development of Baby Buddy as a universal intervention to reduce health inequalities requires robust prospective research that considers effects on the rate and duration of exclusive breastfeeding.

### Acknowledgments

This study was part of the doctoral thesis of LMM, which was funded through the Ho Kong Fung Ling postgraduate scholarship, the University of Sydney, Charles Perkins Centre, and Faculty of Medicine and Health. This manuscript has no direct funding support.

### Authors' Contributions

This project was conducted at the University of Sydney as part of a doctoral thesis (philosophy). LMM and AG contributed to the concept and design of the study. LMM and AG conducted the research and analyzed the data. LMM drafted the first version of the manuscript. AB, NP, and CSEH contributed to writing and editing the manuscript. All authors read and approved the final manuscript.

### Conflicts of Interest

AB is the CEO and founder of Best Beginnings (UK). NP is the evaluation and impact lead at Best Beginnings (UK). All other authors are Australian researchers and declare that they have no competing interests.

#### Multimedia Appendix 1

Reports supplied by Best Beginnings.

[[DOCX File , 20 KB - mhealth\\_v9i4e25668\\_app1.docx](#) ]

#### Multimedia Appendix 2

Using the Behavior Change Wheel (BCW) to analyze breastfeeding video content in the Baby Buddy app.

[[DOCX File , 18 KB - mhealth\\_v9i4e25668\\_app2.docx](#) ]

#### Multimedia Appendix 3

Complete analysis of all breastfeeding items.

[[DOCX File , 20 KB - mhealth\\_v9i4e25668\\_app3.docx](#) ]

#### Multimedia Appendix 4

APEASE (affordability, practicability, effectiveness, affordability, safety, and equity) criteria.

[[DOCX File , 28 KB - mhealth\\_v9i4e25668\\_app4.docx](#) ]

### References

1. Rollins NC, Bhandari N, Hajeebhoy N, Horton S, Lutter CK, Martines JC, et al. Why invest, and what it will take to improve breastfeeding practices? *The Lancet* 2016 Jan 30;387(10017):491-504. [doi: [10.1016/S0140-6736\(15\)01044-2](https://doi.org/10.1016/S0140-6736(15)01044-2)] [Medline: [26869576](https://pubmed.ncbi.nlm.nih.gov/26869576/)]
2. Victora CG, Bahl R, Barros AJD, França GVA, Horton S, Krasevec J, et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *The Lancet* 2016 Jan;387(10017):475-490. [doi: [10.1016/s0140-6736\(15\)01024-7](https://doi.org/10.1016/s0140-6736(15)01024-7)] [Medline: [26869575](https://pubmed.ncbi.nlm.nih.gov/26869575/)]
3. McAndrew FT, Fellows L, Large A, Speed M, Renfrew M. Infant Feeding Survey 2010. UK Data Service. 2012. URL: [https://sp.ukdataservice.ac.uk/doc/7281/mrdoc/pdf/7281\\_ifs-uk-2010\\_report.pdf](https://sp.ukdataservice.ac.uk/doc/7281/mrdoc/pdf/7281_ifs-uk-2010_report.pdf) [accessed 2019-05-14]
4. Horta BL, Loret de Mola C, Victora CG. Long-term consequences of breastfeeding on cholesterol, obesity, systolic blood pressure and type 2 diabetes: a systematic review and meta-analysis. *Acta Paediatr* 2015 Dec 04;104(467):30-37. [doi: [10.1111/apa.13133](https://doi.org/10.1111/apa.13133)] [Medline: [26192560](https://pubmed.ncbi.nlm.nih.gov/26192560/)]
5. Breastfeeding prevalence at 6-8 weeks after birth (Experimental Statistics). Public Health England. 2020. URL: <https://tinyurl.com/e36r9ubv> [accessed 2020-06-20]

6. Breastfeeding in England. Unicef United Kingdom. URL: <https://www.unicef.org.uk/babyfriendly/about/breastfeeding-in-the-uk/breastfeeding-in-england/> [accessed 2019-05-14]
7. Protecting Health and Saving Lives: A Call to Action. Unicef United Kingdom. URL: <https://www.unicef.org.uk/babyfriendly/wp-content/uploads/sites/2/2016/04/Call-to-Action-Unicef-UK-Baby-Friendly-Initiative.pdf> [accessed 2019-05-14]
8. Kennedy R, Mullaney L, Reynolds C, Cawley S, McCartney D, Turner M. Preferences of women for web-based nutritional information in pregnancy. *Public Health* 2017 Feb;143:71-77. [doi: [10.1016/j.puhe.2016.10.028](https://doi.org/10.1016/j.puhe.2016.10.028)] [Medline: [28159029](https://pubmed.ncbi.nlm.nih.gov/28159029/)]
9. Tang K, Gerling K, Chen W, Geurts L. Information and Communication Systems to Tackle Barriers to Breastfeeding: Systematic Search and Review. *J Med Internet Res* 2019 Sep 27;21(9):e13947 [FREE Full text] [doi: [10.2196/13947](https://doi.org/10.2196/13947)] [Medline: [31573903](https://pubmed.ncbi.nlm.nih.gov/31573903/)]
10. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study. *JMIR Mhealth Uhealth* 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: [10.2196/12631](https://doi.org/10.2196/12631)] [Medline: [31215516](https://pubmed.ncbi.nlm.nih.gov/31215516/)]
11. O'Higgins A, Murphy OC, Egan A, Mullaney L, Sheehan S, Turner MJ. The use of digital media by women using the maternity services in a developed country. *Ir Med J* 2014;107(10):313-315. [Medline: [25556255](https://pubmed.ncbi.nlm.nih.gov/25556255/)]
12. Rodger D, Skuse A, Wilmore M, Humphreys S, Dalton J, Flabouris M, et al. Pregnant women's use of information and communications technologies to access pregnancy-related health information in South Australia. *Aust. J. Prim. Health* 2013;19(4):308-312. [doi: [10.1071/py13029](https://doi.org/10.1071/py13029)] [Medline: [24004661](https://pubmed.ncbi.nlm.nih.gov/24004661/)]
13. Hearn L, Miller M, Lester L. Reaching perinatal women online: the Healthy You, Healthy Baby website and app. *J Obes* 2014 Apr 28;2014:573928-573929 [FREE Full text] [doi: [10.1155/2014/573928](https://doi.org/10.1155/2014/573928)] [Medline: [24872891](https://pubmed.ncbi.nlm.nih.gov/24872891/)]
14. Olander EK, Darwin ZJ, Atkinson L, Smith DM, Gardner B. Beyond the 'teachable moment' - A conceptual analysis of women's perinatal behaviour change. *Women Birth* 2016 Jun;29(3):e67-e71. [doi: [10.1016/j.wombi.2015.11.005](https://doi.org/10.1016/j.wombi.2015.11.005)] [Medline: [26626592](https://pubmed.ncbi.nlm.nih.gov/26626592/)]
15. Tackling health inequalities in infant and maternal health outcomes. Department of Health. 2010. URL: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/215869/dh\\_122844.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215869/dh_122844.pdf) [accessed 2019-05-14]
16. Deave T, Ginja S, Goodenough T, Bailey E, Piwek L, Coad J, et al. The Bumps and BaBies Longitudinal Study (BaBBLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app. *Mhealth* 2019 Sep;5(42):42-13 [FREE Full text] [doi: [10.21037/mhealth.2019.08.05](https://doi.org/10.21037/mhealth.2019.08.05)] [Medline: [31620469](https://pubmed.ncbi.nlm.nih.gov/31620469/)]
17. Self Care Project for Parents using Just One Norfolk website and the Baby Buddy app (Final report). Best Beginnings. 2020. URL: <https://tinyurl.com/47vase4z> [accessed 2020-12-05]
18. Michie S, Atkins L, West R. *The Behaviour Change Wheel: A Guide To Designing Interventions*. London, UK: Silverback; 2014.
19. 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](https://doi.org/10.1111/mcn.12768)] [Medline: [30569549](https://pubmed.ncbi.nlm.nih.gov/30569549/)]
20. Gould GS, Bar-Zeev Y, Bovill M, Atkins L, Gruppetta M, Clarke MJ, et al. Designing an implementation intervention with the Behaviour Change Wheel for health provider smoking cessation care for Australian Indigenous pregnant women. *Implement Sci* 2017 Sep 15;12(1):114-114 [FREE Full text] [doi: [10.1186/s13012-017-0645-1](https://doi.org/10.1186/s13012-017-0645-1)] [Medline: [28915815](https://pubmed.ncbi.nlm.nih.gov/28915815/)]
21. Tombor I, Shahab L, Brown J, Crane D, Michie S, West R. Development of SmokeFree Baby: a smoking cessation smartphone app for pregnant smokers. *Transl Behav Med* 2016 Dec;6(4):533-545 [FREE Full text] [doi: [10.1007/s13142-016-0438-0](https://doi.org/10.1007/s13142-016-0438-0)] [Medline: [27699682](https://pubmed.ncbi.nlm.nih.gov/27699682/)]
22. West R, Michie S. *A Guide to Development and Evaluation of Digital Behaviour Change Interventions in Healthcare*. London, UK: Silverback; 2016.
23. Global Strategy for Infant and Young Child Feeding. World Health Organization. 2003. URL: <https://tinyurl.com/nd5pu33k> [accessed 2019-05-14]
24. Focus On: A Proportionate Approach to Priority Populations. Ontario Agency for Health Protection and Promotion. 2015. URL: <https://tinyurl.com/ytna66ej> [accessed 2019-05-14]
25. Marmot M, Bell R. Fair society, healthy lives. *Public Health* 2012 Sep;126 Suppl 1:S4-S10. [doi: [10.1016/j.puhe.2012.05.014](https://doi.org/10.1016/j.puhe.2012.05.014)] [Medline: [22784581](https://pubmed.ncbi.nlm.nih.gov/22784581/)]
26. Field F. The foundation years: preventing poor children becoming poor adults, the report of the Independent Review on Poverty and Life Chances. Digital Education Resource Archive. 2010. URL: <https://dera.ioe.ac.uk/id/eprint/11472> [accessed 2019-05-14]
27. Cooper S. App Pilot Evaluation Report: National in-app data & in-app data from Guys and St Thomas' and Blackpool. Reporting period: 19 November 2014 to 19 May 2015. London, UK: Best Beginnings; 2015.
28. Crossland N, Thomson G, Moran VH. Embedding supportive parenting resources into maternity and early years care pathways: a mixed methods evaluation. *BMC Pregnancy Childbirth* 2019 Jul 22;19(1):253 [FREE Full text] [doi: [10.1186/s12884-019-2388-2](https://doi.org/10.1186/s12884-019-2388-2)] [Medline: [31331285](https://pubmed.ncbi.nlm.nih.gov/31331285/)]



29. Thomson G, Crossland N. Using the behaviour change wheel to explore infant feeding peer support provision; insights from a North West UK evaluation. *Int Breastfeed J* 2019;14:41 [FREE Full text] [doi: [10.1186/s13006-019-0236-7](https://doi.org/10.1186/s13006-019-0236-7)] [Medline: [31548846](https://pubmed.ncbi.nlm.nih.gov/31548846/)]
30. Powell S, Ali Z, Christie S, Apps J, Gooch K. Report on the Evaluation of Baby Buddy M-Health Intervention with a focus on the GSTT Pilot Embedding Site. In. UK: Research Centre for Children, Families and Communities, Canterbury Christ Church University; 2016.
31. Kotter JLC. *Leading Change*. Boston, MA: Harvard Business School Press; 1996.
32. The 8-Step Process For Leading Change. Kotter. URL: <https://www.kotterinc.com/8-steps-process-for-leading-change/> [accessed 2021-04-10]
33. Steinmo S, Fuller C, Stone SP, Michie S. Characterising an implementation intervention in terms of behaviour change techniques and theory: the 'Sepsis Six' clinical care bundle. *Implement Sci* 2015 Aug 08;10:111-111 [FREE Full text] [doi: [10.1186/s13012-015-0300-7](https://doi.org/10.1186/s13012-015-0300-7)] [Medline: [26253306](https://pubmed.ncbi.nlm.nih.gov/26253306/)]
34. Watkins K, Seubert L, Schneider CR, Clifford R. Post hoc evaluation of a common-sense intervention for asthma management in community pharmacy. *BMJ Open* 2016 Nov 18;6(11):e012897 [FREE Full text] [doi: [10.1136/bmjopen-2016-012897](https://doi.org/10.1136/bmjopen-2016-012897)] [Medline: [27864251](https://pubmed.ncbi.nlm.nih.gov/27864251/)]
35. Chiang N, Guo M, Amico KR, Atkins L, Lester RT. Interactive Two-Way mHealth Interventions for Improving Medication Adherence: An Evaluation Using The Behaviour Change Wheel Framework. *JMIR Mhealth Uhealth* 2018 Apr 12;6(4):e87 [FREE Full text] [doi: [10.2196/mhealth.9187](https://doi.org/10.2196/mhealth.9187)] [Medline: [29650504](https://pubmed.ncbi.nlm.nih.gov/29650504/)]

## Abbreviations

**APEASE:** affordability, practicability, effectiveness, affordability, safety, and equity

**BCT:** behavior change technique

**BCTTv1:** Behavior Change Techniques Taxonomy

**BCW:** Behavior Change Wheel

**COM-B:** Capability, Opportunity, and Motivation-Behavior

**DVD:** digital video disc

**OR:** odds ratio

**TDF:** Theoretical Domains Framework

**UK:** United Kingdom

*Edited by L Buis; submitted 10.11.20; peer-reviewed by D Szinay, M Bestek; comments to author 29.11.20; revised version received 09.12.20; accepted 22.03.21; published 15.04.21.*

*Please cite as:*

Musgrave LM, Baum A, Perera N, Homer CSE, Gordon A

*Baby Buddy App for Breastfeeding and Behavior Change: Retrospective Study of the App Using the Behavior Change Wheel*

*JMIR Mhealth Uhealth* 2021;9(4):e25668

URL: <https://mhealth.jmir.org/2021/4/e25668>

doi: [10.2196/25668](https://doi.org/10.2196/25668)

PMID: [33856351](https://pubmed.ncbi.nlm.nih.gov/33856351/)

©Loretta M Musgrave, Alison Baum, Nilushka Perera, Caroline SE Homer, Adrienne Gordon. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 15.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Smartphone Users' Persuasion Knowledge in the Context of Consumer mHealth Apps: Qualitative Study

Eunsin Joo<sup>1</sup>, PhD; Anastasia Kononova<sup>2</sup>, PhD; Shaheen Kanthawala<sup>3</sup>, PhD; Wei Peng<sup>4</sup>, PhD; Shelia Cotten<sup>5,6</sup>, PhD

<sup>1</sup>Department of Public Relations and Advertising, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, China

<sup>2</sup>Department of Advertising and Public Relations, Michigan State University, East Lansing, MI, United States

<sup>3</sup>Department of Journalism and Creative Media, University of Alabama, Tuscaloosa, AL, United States

<sup>4</sup>Department of Media and Information, Michigan State University, East Lansing, MI, United States

<sup>5</sup>Department of Sociology, Anthropology, and Criminal Justice, Clemson University, Clemson, SC, United States

<sup>6</sup>Department of Communication, Clemson University, Clemson, SC, United States

**Corresponding Author:**

Eunsin Joo, PhD

Department of Public Relations and Advertising

Beijing Normal University-Hong Kong Baptist University United International College

2000 Jintong Road

Tangjiawan

Zhuhai, 519087

China

Phone: 82 1086813554

Email: [eunsinjoo@uic.edu.hk](mailto:eunsinjoo@uic.edu.hk)

## Abstract

**Background:** Persuasion knowledge, commonly referred to as advertising literacy, is a cognitive dimension that embraces recognition of advertising, its source and audience, and understanding of advertisers' persuasive and selling intents as well as tactics. There is little understanding of users' awareness of organizations that develop or sponsor mobile health (mHealth) apps, especially in light of personal data privacy. Persuasion knowledge or recognition of a supporting organization's presence, characteristics, competencies, intents, and persuasion tactics are crucial to investigate because app users have the right to know about entities that support apps and make informed decisions about app usage. The abundance of free consumer mHealth apps, especially those in the area of fitness, often makes it difficult for users to identify apps' dual purposes, which may be related to not only helping the public manage health but also promoting the supporting organization itself and collecting users' information for further consumer targeting by third parties.

**Objective:** This study aims to investigate smartphone users' awareness of mHealth apps' affiliations with 3 different types of supporting organizations (commercial, government, and nonprofit); differences in users' persuasion knowledge and mHealth app quality and credibility evaluations related to each of the 3 organization types; and users' coping mechanisms for dealing with personal information management within consumer mHealth apps.

**Methods:** In-depth semistructured interviews were conducted with 25 smartphone users from a local community in midwestern United States. Interviews were thematically analyzed using inductive and deductive approaches.

**Results:** Participants indicated that their awareness of and interest in mHealth app-supporting organizations were secondary to the app's health management functions. After being probed, participants showed a high level of persuasion knowledge regarding the types of app-supporting organizations and their promotional intents. They thought that commercial companies sponsored mHealth apps mostly as entertainment tools, whereas noncommercial entities sponsored mHealth apps for users' education. They assigned self-promotional motives to commercial organizations; however, they associated commercial mHealth apps with good quality and functioning. Noncommercial entities were perceived as more credible. Participants were concerned about losing control over personal information within mHealth apps supported by different organizations. They used alternative digital identities to protect themselves from privacy invasion and advertising spam. They were willing to trade some personal information for high-quality commercial mHealth apps. There was a sense of fatalism in discussing privacy risks linked to mHealth app usage, and some participants did not perceive the risks to be serious.

**Conclusions:** The discussion of and recommendations for the safe and ethical use of mHealth apps associated with organizations' promotional strategies and personal data protection are provided to ensure users' awareness of and enhanced control over digitalized personal information flows. The theoretical implications are discussed in the context of the Persuasion Knowledge Model and dual-processing theories.

(*JMIR Mhealth Uhealth* 2021;9(4):e16518) doi:[10.2196/16518](https://doi.org/10.2196/16518)

## KEYWORDS

mHealth app; personal health information sharing; mobile phone; mobile promotion strategy; persuasion knowledge

## Introduction

### Background

In today's mobile-driven era, mobile apps have a significant impact on users' healthy lifestyles [1]. Using mobile app services can help individuals manage chronic diseases and healthy lifestyles as well as fight bad habits, such as smoking [2,3]. Commercial, governmental, and nonprofit organizations support mobile health (mHealth) apps in diverse formats, and many do it with a dual purpose: to help improve public health and to promote their organizations [4]. For example, Under Armour sponsored and later purchased MyFitnessPal, a leading free app for achieving and maintaining health and fitness goals [5,6]. The Baby Center portal, formerly owned by Johnson & Johnson, and a corresponding pregnancy tracking app reached 45 million users in 2016 [7]. mHealth app support occurs not only in commercial sectors but also in public sectors, including nonprofit and government agencies. A US-based Centers for Disease Control and Prevention's new parenting app, for example, offers customized services for parents to track their children's developmental milestones [8]. The nonprofit organization the American Red Cross launched a blood donor app that enables the organization to communicate with its donors and offers a rewards program for users to increase the frequency of blood donations [9].

Despite the benefits to app users and app-supporting organizations, the growing consumer mHealth app market raises concerns over users' awareness of app-supporting organizations and personal information privacy because personal data can be shared via mobile phones and wireless networks [2,10-13]. The mobile app market is dominated by free-to-use consumer apps, with more than 90% available without charge in the major app stores (ie, Apple App Store and Google Play) [14]. Instead of paying for such apps with money, mobile users may be asked to give away bits of personal information needed for an app to function. Personal information (eg, age, gender, and email), including basic health information (eg, weight and height), can be shared with unauthorized third parties via mobile phones and users do not easily recognize the magnitude of such information sharing [15-20].

Access to users' personal information becomes a concern because mHealth apps' affiliations with supporting organizations, especially those in the commercial sector, are often obscure [7]. This adds complexity to users' interpretations of organizations' motivations to support these apps. On the one hand, such support is driven by the intent to help users manage their health and help others. On the other hand, supporting organizations gain direct access to consumers' personal data

and may use these data for commercial, marketing, and other self-serving organization-related purposes (eg, fundraising or customized advertising). This emphasizes the value of studying users' critical assessment of mHealth apps associated with different types of organizations, users' knowledge of organization types and intentions to support mHealth apps, and users' strategies to cope with privacy-related risks when they are asked to share personal information via smartphones.

This study contributes to the existing literature on persuasion knowledge and mHealth apps that serve the dual purpose of improving public health and promoting supporting organizations. Previous mHealth research has examined app selection process in app stores and issues related to potential privacy risks of personal information collection and management [3,10,21]. However, there is little understanding of whether and how smartphone users react to cues about organizations that support mHealth apps and how they understand and negotiate the duality of organizations' motivations (ie, users' health management vs organization self-promotion). Furthermore, few studies have explored differences in users' perceptions of mHealth app quality and credibility and willingness to share personal information across organization types: commercial, government, and nonprofit, in light of organizations' dual motivations to support such apps. To address these gaps, this study applies the Persuasion Knowledge Model (PKM), dual-processing theories, and information privacy literature to explore how mHealth apps associated with commercial, government, and nonprofit organizations influence smartphone users' understanding of mHealth apps. It investigates mHealth app use not only as a health management tool but also as a promotional tactic. The study provides useful insights about mHealth app cues associated with commercial and noncommercial organizations that can be used to efficiently communicate mHealth app affiliation to users and enhance their critical assessment of this health-related technology. This study also provides suggestions for government and nonprofit organizations to offer effective and engaging mHealth app services that ensure individual autonomy in protecting personal data.

### mHealth App Market

The mHealth app market was valued at approximately US \$12.4 billion in 2018 and is expected to expand at a compound annual growth rate of 44.7% from 2019 to 2026 [22]. The World Health Organization broadly refers to mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices" [23]. Some mHealth apps are provided by medical organizations to communicate personal health information (PHI; eg, medical test results, prescriptions,

and diagnosis) to patients. Such apps fall under the category of personal health records. They provide a single space for patients to access their own health records and simplify patient-provider communication [24]. Alternatively, many mHealth apps can be downloaded from app stores without providers' involvement in the user's initiative to maintain a healthy lifestyle or to enhance the health of others [25]. These apps are often developed or owned by companies with aligned stakes, such as MyFitnessPal, or standalone health-related apps, such as Flo, the period tracking app. The information provided to the app comes from the user based on their own knowledge of their health and data tracked based on app usage. Such apps have no direct connection to health professionals.

Today, almost 60% of US smartphone users have downloaded at least one health-related app on their mobile devices [26], with the exercise and weight loss app category being the most popular [3,4]. It is common for users to provide personal information, including basic health information (eg, height, weight, BMI, physical activity levels, calorie and water intake, pregnancy status, sleeping patterns, mood, period, and sexual activity), for customization by direct input or connecting to sensors and wearable technologies [4,5,10]. This information is typically less protected, especially if mHealth apps are free to use. Sharing users' data with digital marketers within and outside the organization becomes a source of the sustainability of these apps [2,18]. This study explores smartphone users' perceptions of these consumer mHealth apps, with the focus on persuasion knowledge and privacy concerns.

### PKM and Promotional mHealth Apps

Persuasion knowledge is important for users with respect to mHealth apps because they are created for users' health management and organizations' promotional purposes. Conceptually, persuasion knowledge, commonly referred to as advertising literacy, is a cognitive dimension that embraces recognition of advertising, its source and audience, and understanding of advertisers' persuasive and selling intents as well as tactics [27]. Conceptual persuasion knowledge is different from evaluative persuasion knowledge that deals with consumers' affective evaluations of advertising. According to the PKM, people develop and use their knowledge derived from previous experience, education, and socialization to recognize, interpret, evaluate, and respond to persuasion attempts, such as advertising [28]. As a reaction to persuasion (promotion) attempts, media users choose and execute persuasion coping behaviors that they perceive as effective and appropriate. Such behaviors may be positive for a brand or organization that promotes itself using an mHealth app (eg, buying and telling others about the organization or app) and may be negative (eg, deleting the app and boycotting the organization) [29]. In the contexts of brand-related apps, the more favorable users feel about a brand-supported technology, the more receptive they are to its claims or content, including the identification of a supporting organization [30,31].

The PKM includes 3 belief structures [28]. The first structure refers to persuasion knowledge itself, where consumers are aware of and understand actors in the self-promotion persuasion process, persuaders' intentions, and tactics used to persuade,

among others. In relation to promotional mHealth apps, persuasion knowledge may refer to app developers and sponsors who put the app on the market, understanding the reasons for app support and perceptions of the app as a promotional tool to increase sales, donations, public awareness, and other desirable outcomes. The second structure is related to the perceptions of the persuasion agent (eg, advertiser), including the agent's traits, competencies, and goals. In the context of this study, we discuss 3 types of persuaders: commercial, governmental, and nonprofit entities. Understanding the nature of each persuader, its resource base, and mission constitutes beliefs about the agent. The third structure is associated with beliefs about the topic of persuasion (eg, product, service, social cause, or candidate). It could be, for example, related to a specific health issue that an app focuses on (eg, fitness and diabetes) and the digital app itself (eg, MyFitnessPal).

When consumers' persuasion knowledge levels are high, they are able to understand the self-promotional intent and, as a result, doubt the altruistic intentions of an organization and activate cognitive defenses against persuasion [21,32,33]. Consumers with high persuasion knowledge are more likely to be skeptical of self-promotion persuasive communication and resistant to persuasive advertising messages or sponsored products [29,34]. Persuasion knowledge plays a significant role in evaluating subtle (vs prominent) digital internet-based advertising formats [29,34]. For example, users with a higher level of persuasion knowledge may be more likely to recognize a supporting organization when they examine an mHealth app (especially if such an app is offered for free) and form their own beliefs about the organizations' intentions, be they for the public good (eg, social responsibility and establishing healthy lifestyles) or self-service (eg, profit seeking, data collection and sharing with third parties, and social control).

When evaluating an organization's support of consumer technology, such as mHealth apps, and making usage decisions, users may negotiate between organizations' self-promoting and user-oriented motives differently across different types of app providers, especially if they are asked to give away personal information to download and use the app [12,16,35,36]. Previous studies have indicated that nonprofit sponsors of information and communication technologies receive more favorable evaluations than their for-profit counterparts. This is reflected in the positive attitudes and attribution of less egoistic motives to nonprofit entities [36,37]. Although consumers may have difficulty inferring selfish motives in cases of prosocial persuasion attempts, such as antidrunk driving campaigns (or, in the case of this study, public health), they would still trust nonprofit or government agencies as campaign sponsors more than corporate organizations [37]. We suggest and further explore the type of organization associated with different degrees of negotiation between self-serving and public service motives when smartphone users decide to download an mHealth app and share personal information within it. This study aims to examine the nature of persuasion knowledge and mHealth app evaluations associated with 3 types of app-supporting organizations (ie, commercial, governmental, and nonprofit) and to determine if willingness to share information within mHealth apps differs by organization type.

## Sharing Personal Information via mHealth Apps

Many smartphone apps require users to give permission to access personal data (eg, social media data and contact lists) and phone functions (eg, camera and speakers), and such permission is very easy to obtain (eg, clicking “I Agree”). As a result, issues related to sharing personal information with the app are emerging as an important research area for health practitioners and policy makers [10-12,19,20]. Personal information obtained via the apps can be shared within the supporting organization itself as well as with unknown third parties, such as marketers, without explicitly notifying app users [15-20]. At the initial level, users might not realize that the app they use is affiliated with an organization. For example, formerly Johnson & Johnson’s Baby Center portal and app used to have only subtle cues (eg, pop-up advertisements) about such affiliation. Furthermore, users may not always realize that the personal information they provide is used within and outside the company for marketing purposes. This may facilitate the risks from emotional distress to financial discrimination. For example, period and pregnancy tracking consumer apps are popular not only with female smartphone users but also with companies, such as Johnson & Johnson, which target this demographic with offers of women- and parenthood-related health products. Users share sensitive health information with such apps, including patterns of sexual activity, number of pregnancies and miscarriages, and gestational age of unborn children. Mere awareness of such data being shared for marketing purposes may create psychological discomfort. Moreover, other companies may be interested in accessing this target group’s data to promote relevant products (eg, cars and realtor services). Although the data are deidentified and aggregated when moving through a complex analytics process, they help profile individuals and assess their qualifications for life insurance, mortgages, and loans [38-40].

The classical definition of privacy is the freedom to protect oneself from exposure to or intrusion by others [11,41]. Privacy is an important requirement in the health care domain that deals with the challenges of maintaining PHI about one’s condition and health history confidential while sharing them with authorized medical parties and caregivers, guardians, and family members [42]. Although strict rules apply to protecting one’s PHI [2], policies to manage one’s personal information (including basic health information) provided to consumer mHealth apps are less clear.

In the digital age, this is discussed in the context of having autonomy. Autonomy refers to an individual’s right to control the environment in which they live in and make rational decisions [43]. An mHealth user, for example, can decide to download or delete the app at any time; thus, they have control over using it. However, can they make such decisions about personal and health data collected via this app? Can they recognize persuasion attempts when they see marketing messages tailored to them using sophisticated technologies? Privacy is described as a tool that fosters and encourages autonomy. Becker [43] argues that omnipresent digital technologies that allow constant surveillance put at risk not only privacy—when an individual is *observed* on the web and *analyzed* via algorithms—but also autonomy where the

individual loses the power to make informed decisions about their personal information flows (especially when information is deidentified and aggregated) as well as make independent decisions about customized marketing messages.

Some users avoid downloading mHealth apps because of privacy concerns and potential risks related to the collection of personal identifiable information [3,10,35]. Users also consider advertising messages and commercial identification (eg, brand logo) as negative cues when making credibility judgments of sponsored websites [36,37]. According to the literature on dual-processing theories and decision-making processes, users tend to rely on heuristic cues to simplify the selection process if they face cognitive limitations because of a deluge of information [11,35,44]. Such heuristic-based decision-making processes may promote automatic app judgment and selection that does not require much effort for a thorough app evaluation [11,45]. Users, for example, rarely read app privacy policies; that is, they rarely engage in systematic, elaborate, and effortful information processing before downloading an app unless they are first informed about negative consequences of sharing personal information [11,18,19]. Instead, they relied on app visual cues to assess it [45]. Many do not worry, are unaware of data sharing with third parties, or cannot imagine the large scope of such sharing [13,19].

Sources of digital content (eg, developer, owner, or sponsor in the context of mHealth apps), if visible and recognizable, may serve as a powerful cue to activate heuristics that guide the evaluation of the digital content [35-37]. These heuristics may differ according to the type of source. In this study, we suggest that smartphone users have different mental representations of commercial, governmental, and nonprofit organizations that support mHealth apps. Such differences manifest themselves in unique perceptions of each organization type, including its characteristics, capacities, and self-promotion versus public service motivations, differences in assessing the quality and credibility of mHealth apps associated with different types of supporting organizations, and different levels of willingness to share PHI with supported mHealth apps. Understanding these differences will inform the development of transparent mHealth technologies that effectively communicate information about technology affiliation with an organization and equip users with strategies to protect their personal information when engaged with technology use.

## Research Questions

Applying the PKM [28], dual-processing models [11,44,46], and the concepts of information privacy discussed earlier, we ask the following research questions (RQs):

- RQ1: What are study participants’ levels of awareness of and interest in knowing an mHealth app affiliation with a supporting organization?
- RQ2: What are the differences in the nature of participants’ persuasion knowledge (eg, recognition of target and agent and understanding of agent’s characteristics and capacities, self-promotion, and public service purposes) across commercial, governmental, and nonprofit organizations that support mHealth apps?

- RQ3: In what ways do participants' evaluations of mHealth apps' quality and credibility differ by the type of supporting mHealth app organization?
- RQ4: What are the differences in coping mechanisms, if any, that participants implement when they are informed about mHealth app support by commercial, governmental, and nonprofit organizations, especially in light of sharing personal information within such apps?

## Methods

### Recruitment

In-depth semistructured interviews were conducted at a large university in the midwestern region of the United States. The study protocols were approved by the institutional review board before data collection began. Local community residents (N=25) were recruited through a web-based recruitment pool. Each participant received US \$15 for their participation in the research. Data collection was stopped at a sample size of 25 because of response saturation. Specifically, the last participants interviewed confirmed the responses from earlier interviews and did not provide additional novel insights [47,48].

### Procedure

The average time for each interview was 52 minutes, ranging from 36 to 92 minutes. All interviews were recorded using digital recording options available on the interviewers' smartphones. A total of 3 researchers trained 2 student interviewers and oversaw the data collection process. Participants were informed about being recorded in the consent document before they started the interview. All participants had the option of withdrawing from the study or refusing to answer any question; none of them did. The full interview guide is presented in [Multimedia Appendix 1](#). Participants were asked about their perceptions of *the government*, *commercial*, and *nonprofit* apps. Examples were provided per the participants' request. Most of our participants were familiar with the concept of mHealth apps and did not need an explanation as to what health apps were.

### Data Analysis

Anonymized interview audio files were transcribed using a web-based transcription service. Data coding was continued with 4 coders (coauthors). As per standards of qualitative methodology and, specifically, thematic analysis [49], the coders were open in their approaches to raw data analysis and flexible in revising transcript interpretations. NVivo, a qualitative data analysis computer software, was used to organize the codes and corresponding data. To assess coders' agreement with the generated codes, 3 identical transcripts were independently analyzed. Coders met multiple times to discuss emerging codes. The coding tactics were refined iteratively in each meeting. First, coders discussed each code they identified, both on descriptive and interpretative levels [50], and unified names and definitions of codes that they agreed upon. After the initial rules of coding were established based on multiple readings of the same transcripts and each code received a clear definition, 2 coders coded the remainder of the transcripts. A total of 139 codes and subcodes were developed with 538 references (Table

S1 in [Multimedia Appendix 2](#)). The codes and subcodes were then analyzed to inductively derive major themes [51].

We chose to base our coding on empirical data collected more than on previous literature (ie, inductive approach) because of the centrality of organization type to the conversation between participants and interviewers. As little work, to our knowledge, has been done about the perceptions of mHealth apps supported by government, commercial, and nonprofit organizations, we did not impose a strict top-down structure (deductive approach) on the coding rubric. A deductive approach was used after the themes were derived to organize our reporting of the findings in accordance with the theoretical frameworks used and the RQs examined.

To ensure the trustworthiness and consistency of the findings, we used a number of procedures [48,49,52]. These included team-based instrument development designed to achieve the study's objectives, extensive training of interviewers and coders, data collection oversight, using multiple coders to work with transcripts, developing a coding rubric and establishing consistent analysis routines (regular discussions of code and theme interpretations), applying logic to assess code relevance and irrelevance, and supporting themes with quotations from participants.

## Results

### Overview

The participants' demographic descriptions are provided in Table S2 ([Multimedia Appendix 3](#)). The findings provide descriptive information about the participants' use of smartphones, apps, and mHealth apps and report themes derived through the analysis ([Multimedia Appendix 2](#)).

### Mobile Phone and App Use: Descriptive Information

More than half of the interviewees had an iPhone (13/25, 52%), 6 had a Samsung smartphone, and 6 had other types of smartphones. The average length of smartphone use in the sample was approximately 3 years. On average, interviewees had 29-32 apps on their phones. The most used apps were social media apps (Facebook, Instagram, Snapchat, etc). Other frequently used app categories included email, messaging, video chatting (eg, Skype), utility (eg, weather and maps), game, shopping, banking, entertainment streaming (eg, Netflix and Spotify), and health and fitness (eg, MyFitnessPal) apps. In total, 9 of the 25 (36%) participants reported that they had at least one health-related app on their smartphones. Most health-related apps mentioned by the interviewees were related to maintaining healthy lifestyles (eg, step counter, calorie counter, running tracker, healthy eating, and meditation), and 2 mentioned disease or disorder management apps (eg, attention deficit hyperactivity disorder).

### RQ1 and RQ2 Findings

RQ1 asked about study participants' levels of awareness of and interest in knowing an mHealth app affiliation with supporting organizations.

### **Recognition of mHealth App Source Is Secondary to the App's Health Management Functions**

Participants expressed little awareness of and curiosity about promotional app support before we probed them with questions related to the study's RQs. Most participants said that they rarely paid attention to information about mHealth app developers, sources, and sponsors. For them, mHealth apps' ease of use, information quality, relevance, and functionality or utility were more important than an entity supporting the apps:

*If it records what I want, I do not care who designed it. If it's easy to use and if it has all the information I need, then that's why I would do it.* [Participant 9, female, aged 65 years]

Many participants elaborated that any mobile app was easy to delete; thus, apps did not pose any danger and did not lead to users' personal data breaches. For example, one participant (participant 14, female, aged 29 years) indicated that the use of promotional mHealth apps was not a *serious* issue. Thus, at the initial level of discussing promotional mHealth apps, users were consumed with the health management function of these apps and did not consider additional, self-promoting motivations that could drive the organization's app support.

RQ2 asked about the differences in the nature of participants' persuasion knowledge across commercial, governmental, and nonprofit organizations that support mHealth apps. After being probed, most interviewees indicated that they possessed persuasion knowledge of app-supporting organizations. They perceived themselves and similar groups of consumers as well as society as a whole to be the targets of persuasive attempts initiated mostly by commercial companies.

### **From Commercial Entertainment to Noncommercial Information**

Discussing the characteristics and capacities of organization types, participants were more likely to assign *general information* function to government agencies. Overall, increasing public awareness of health problems and educating people about them was the overarching goal of governmental and nonprofit agencies and associations. As for commercial companies, their products and services, while associated with the best quality, were mostly perceived as *fun* and *entertaining*:

*Bono's probably on there [commercially supported mHealth apps] talking or something. They've probably got a quote from Jay-Z on there.* [Participant 23, male, aged 25 years]

Government agencies were perceived as being research driven and resourceful in terms of the health information available. As the government's role in supporting mHealth apps was mostly discussed as being a general, broad information provider, its mHealth apps were perceived as secondary to the web resources that provided a great deal of credible medical information to participants. It was easier for participants to access these web resources on their computers, rather than via mHealth apps:

*If I wanted to search Metformin, diabetic drug, then everything that's on the web on that...* [Participant 9, female, aged 65 years]

*I'll search stuff like that online, on the Internet, but I don't need an app for it...* [Participant 24, female, aged 24 years]

Narrow specialization, knowledge of one problematic health area, was the prerogative of nonprofit organizations:

*Their goal is to provide the health information or the healthcare information that is needed by the user without focusing on other areas that the person doesn't need. If I have to find out about cancer, I'm [going to] go to the American Cancer Society app, I'm not [going to] go to the one about diabetes.* [Participant 10, male, aged 65 years]

### **Self-Promotion and Public Service Motivations Behind mHealth Apps**

Participants recognized strong self-promotion intentions of for-profit organizations to support mHealth apps, such as selling products, advertising a company, and building brand image. Some explained that commercial companies used apps to misinform users and collect consumers' *opinions*, record personal information, and track app usage data. Only a few attributed commercial app support to corporate social responsibility oriented to public service:

*That's like an advertising platform for them. So, if you were to get an app to count calories while you were running, that app's created by Nike, at some point it would try to sell you some kind of Nike shoes. So yeah, it's an advertising platform for commercial organizations besides, of course, getting some consumer goodwill.* [Participant 11, male, aged 40 years]

Although participants could clearly distinguish between the intentions of commercial and noncommercial organizations, there was little understanding of the differences between the government and nonprofit sectors. Participants expressed great concern about paying *my money* for commercial products and services but they were much less worried about how taxpayers' and donors' money was spent by government and nonprofit organizations. They perceived nonprofit and government agencies to be more careful and accountable in spending, especially that these agencies, according to the participants, had a clear goal of improving public health by helping people via mHealth apps. Thus, public-serving motivation was more pronounced in participants' perceptions of noncommercial entities. Only a few participants mentioned that nonprofit and government organizations developed mHealth apps for self-profiting reasons: to collect personal information and to fundraise for a cause.

The topic of trust in an organization's public service motivation was related to the discussion of congruency between an app-supporting organization and the health issue or cause. Interviewees expressed distrust in some commercial organizations that would support an irrelevant health issue. One participant called it *hypocritical*. Congruency was identified not only on the level of the cause but also on the level of the organization's general mission. Interviewees found that it was more relevant for nonprofit organizations to support health apps

because it was consistent with their goals of making people healthier and contributing to the overall social good:

*I would always pick the government organization or the non profit organization, because I feel they are doing it for the public good, because they don't have anything to gain by it, as opposed to somebody [commercial companies] who is doing it...to make money off of it. [Participant 20, female, aged 54 years]*

### RQ3 and RQ4 Findings

RQ3 asked how different types of supporting mHealth app organizations influenced participants' evaluations of mHealth apps' quality and credibility.

#### Quality Does Not Mean Credibility

The discussion of advertising literacy and mHealth app evaluations centered on 2 aspects of app evaluation: perceived app quality and perceived app credibility. We left the definitions of these constructs open so that participants interpreted them according to their definitions of *quality* and *credibility*. Overwhelmingly, participants discussed mHealth app *quality* as related to the apps' looks, usability, and functionality while they talked about *credibility* as related to trust and intentions of supporting organizations:

*I expect [commercially supported app] to look fancier and brighter. I will say the quality to be better as in its functionality, it shouldn't have any slowness when you download it, type thing. Not actually having more options, just its functionality should be better from big companies...I don't put much credibility into their things [commercial companies] because a lot of things they do are for profit...I like non profits, I would do something from American Cancer Society probably easier than I would accept something from Coca-Cola. [Participant 13, female, aged 23 years]*

Perceptions of organizational resourcefulness were directly linked to mHealth quality judgments. A distinguishing feature of quality was related to *dumping more money* into app development. Commercial organizations were perceived as more willing to invest in better quality apps compared with *less rich* nonprofit organizations or government agencies. There was no expectation for publicly (not commercially) supported apps to look appealing:

*They [nonprofit and government organizations] might get the job done, they might capture the information, but it's not going to keep my interest, or be visually appealing, or have the functionality that I need. And so, I still wouldn't pay for it though, but I may have to look at a [commercially] sponsored app. [Participant 2, female, aged 48 years]*

A prominent perceived characteristic of government-backed apps was related to providing scientific, research-supported health information and help for people:

*I would give credibility to the Health and Human Services or a Center for Disease Control. It all depends on the particular information I'm looking for, whether it's about preventing the spread of a*

*disease, or if it's developing good practices for personal health care. The government has a place for its information. [Participant 10, male, 65 years]*

Another valuable characteristic of government-based mHealth apps was security. With this came trust in government-supported digital products:

*I would trust more in a governmental app in this case, just because they have more rules about security and it's probably harder to get access to governmental data. [Participant 21, anonymous]*

Overall, participants described commercial organizations as less credible because self-promotion and profiting were perceived as the obvious drivers of mHealth app support. Nonprofit organizations and government agencies were described as more trustworthy than their commercial counterparts. However, some participants mentioned the complexity of indicating true motives in relation to all 3 organization types because they supported mHealth apps not only *to make people healthier* but also for publicity and control.

Despite a clear understanding of commercial organizations' intentions, participants leaned toward using commercially supported mHealth apps, as they were better in appearance and features. In other words, they valued app quality more than information credibility. Few participants were willing to *sacrifice* visual appeal and functionality for more *ethically* supported and credible mHealth apps from noncommercial sources.

RQ4 asked about the differences in coping mechanisms, including the strategies to protect personal information shared electronically, that participants implemented when they were informed about mHealth app support by commercial, governmental, and nonprofit organizations.

#### Trade-Offs

Most participants, while having privacy concerns, did not have the habit of reading privacy terms and conditions before downloading and using mHealth apps. Instead, they used social and heuristic cues, such as star ratings and user reviews, to make downloading decisions. Only one participant claimed that he had read app privacy policies.

Many agreed to *pay* for mHealth apps use with their personal information and increased risk of this information being used for purposes not related to app functioning:

*As long as it is not my social security number and my credit card number, I tend to give that information [to app supporting organizations]. It might be a win-win. I mean it might also help me manage my health but of course the commercial for-profit companies are always interested in my opinion, and always interested in making sure that you buy their product or are aware of their product. [Participant 2, female, aged 48 years]*

When participants had to share personal information, they indicated that they preferred to share less information than more information and general information over specific medical information:



*I would be willing to give more general types of information, as opposed to very specific information, like, "Do you take 150 micrograms of Lexo, whatever, a day?" I would not [want to] say that's what I do, but I would say, I'm on a thyroid medication.*  
[Participant 20, female, aged 54 years]

Furthermore, participants tended to agree to share greater amounts of personal information with nonprofit or governmental organizations than commercial organizations to use mHealth apps because they believed that nonprofit and government agencies used such information for public health and research:

*I'm going to probably choose something that is either government or non profit related before I would choose something that's for-profit before, just because I am assuming that for-profit it's probably [going to] want more of my information to try to get me to purchase their products.* [Participant 8, female, aged 36 years]

Many participants were open to providing some personal information to download and use mHealth apps:

*It doesn't bother me, I think I'm not a superstar.*  
[Participant 14, female, aged 29 years]

They shared basic personal and health information, but it depended on how much information the app needed them to share and whether this information was relevant to the purposes of the app. Demographic information that participants were willing to share commonly included name, email, and date of birth. Interviewees shared health information if they were relevant to basic app functioning. There were also some differences in the type of information that could be shared with apps provided by companies, nonprofits, or government agencies:

*If I have to put in my personal information, it depends on how much personal information. In order to use the [mHealth] app, I have to tell them [app developers and supporters] how old I am and how much I weigh and how tall I am and those kind of [things].*  
[Participant 4, female, aged 34 years]

*I avoid trying to give out my postal address because most times in commercial, it's just gonna give you junk mail...My weight, my height are fine...Depending who the non profit organization, I very well may give more information than a commercial...I wouldn't give them like the lab paperwork, but I would be more interested in saying, "Hey, this is what may or may not be going on with me and this is what the kind of information I'm willing to give you about myself." So, I feel that non profits can get a little bit more information because they're not in it for the money.*  
[Participant 7, male, aged 28 years]

In addition, some participants suggested that signing in with an existing account on a social networking site, such as Facebook, was an easier way to download and use an mHealth app rather than creating a new, app-specific log-in. Participants did not see this type of information sharing as a threat to their privacy,

as they did not consider personal information shared on social media to be private:

*Yeah, I do that, I'll connect, say that way they can, I guess I give them some of the information. Because if I'm posting on Facebook, I feel that's generally public information amongst everyone.* [Participant 7, male, aged 28 years]

Convenience was the key reason for using social media sign-ins instead of creating an app-specific password. It was, as few expressed, a good way to save time and energy put into registering and accessing a new mHealth service.

### ***It Is Culture and It Is Normal***

Some interviewees expressed a feeling of fatalism and helplessness when it came to using promotional apps and sharing personal data with the supporting organizations and third parties. One participant, for example, understood the scope of the issue; yet, the issue was systemic and too big for one individual to combat:

*I think I just assume that anything that I input into an app or put online, it could be seen by someone, or by a company, or whatever. Someone could be tracking it and so, I don't think I necessarily put too much thought into privacy, other than with the Mood Tracker app that I was looking at where you could choose to pay more and have it be private so it looked like other users won't be able to see your data, but the company can still see your data. I guess I feel like there's not really options of...No matter what app you have, somebody's [going to] be able to see it, somebody's [going to] be able to track your data and so it's just the risk you take I guess in doing that.*  
[Participant 12, female, aged 24 years]

Another participant, an international resident of the local community, looked at this overpowering phenomenon as a consequence of the culture of transparency. As it was part of the culture, it was normalized to the participant's mind:

*I think I'm already kind of used to this culture here that, because if you go to see dentist, you're, also when you're providing any information about what happened to you. That's very important, like you watch, kind of, you're alert to what kind of things. That's very normal, so you need to provide that.*  
[Participant 14, female, aged 29 years]

### ***Defenses Against mHealth Persuasion Attempts***

Some participants indicated they were less willing or unwilling to share any type of personal records. One reason was the potential to raise security risks related to identity theft and financial losses. Another reason was linked to privacy invasion or sharing personal information that participants preferred to keep private or have control over:

*I'm not a huge fan of the intrusion via apps [...] I guess I'm not comfortable with how much companies can learn about me without my permission.*  
[Participant 6, female, aged 23 years]

Advertisements, junk mail, and email from commercial companies contributed to another reason the interviewees were averse to providing personal information; they were worried about app providers using their information for commercial purposes rather than for the needs of the app. Commercial purposes were associated with traditional print and digital advertising, such as receiving direct mail and promotional emails (spam):

*Yeah, I avoid trying to give out my postal address because most times in commercial, it's just [going to] give you junk mail. [Participant 7, male, aged 28 years]*

*I would be less likely to put that information to a commercial app than to a non profit app, because I would be wondering what they were [going to] do with that...It's a good way of cheap advertising. [Participant 20, female, aged 54 years]*

Several participants admitted that they used alternative personal information, such as secondary email addresses or fake names, to sign in to mHealth apps and thus avoid potential privacy invasion by third parties. This finding suggests that technology users seek creative ways to deal with privacy breaching threats that do not require extensive mental work to go over pages of hard-to-read terms of use:

*I would use a name and date of birth, but it's [going to] be fake for me almost no matter what website I sign up for. When it's just some app or some website I'm not as apt to do that, but I'll give them a fake name. I don't really like the date of birth, but I know that sometimes you have to verify age. I wouldn't give an address really for any app. [Participant 18, male, aged 23 years]*

## Discussion

### Principal Findings

The results of the study showed that participants were admitted to not paying much attention to mHealth app-supporting entities. An interesting strategy identified through our research is related to the tricks that our interviewees used to *play the system*. While understanding persuasive intents and having concerns about personal information privacy, participants did not engage in reading terms of use when downloading mHealth apps. Instead, they tried to protect themselves by using fake names and email or social media accounts that were designated to receive *junk* advertising messages.

When probed, the participants recognized the intentions behind commercial support. Most participants possessed high levels of persuasion knowledge. They were more likely to assign selfish than altruistic motives to commercial companies that support mHealth apps. Participants believed that nonprofit and governmental organizations spent their funds with greater accountability and supported health apps to protect public health. Although they distinguished between the intentions of commercial and noncommercial organizations, there was a lack of clear distinction between government and nonprofit agencies.

Many participants agreed to provide personal information to download and use mHealth apps. However, some participants expressed concerns about doing it because of potential risks related to invasion of privacy, data security, junk mail, and possible misuse of information by a third party. The discussion of *annoying* advertising messages that target app users to sell products was much more prevalent than the concern of sharing personal information with unknown parties without user awareness and consent. Participants did not reveal in-depth knowledge of the personal information sharing process in the digital sphere (eg, the use of artificial intelligence in advertising, algorithmic ad delivery based on digital data clusters, and programmatic buying). This leads to an important conclusion that the consequence of personal information sharing, including basic health information, shall be experienced in a relatively direct way, where individuals make straightforward associations between information sharing and receiving unpleasant promotional messages.

When the discussion touched on more abstract topics related to privacy invasion and required an understanding of complex processes that support the practices of data sharing through mobile and wireless networking technologies, participants were less likely to be concerned. Some participants, however, were less likely or unlikely to share personal information with consumer mHealth apps, regardless of the type of source organization. These participants had the general idea of abstract dangers related to privacy breaching and used the strategy of *being on the safe side* by distrusting any mobile app data sharing requests. Some interviewees, even if they decided to provide personal information, as it is a common element of the contemporary app use culture, were still concerned about the misuse of their personal information that would serve organizations' commercial interests instead of being used exclusively for the app's direct purposes. This suggests that it is possible that participants might not entirely understand the mechanism of misuse but they are at least roughly aware of what is happening while using data sharing apps. These findings suggest that some foundational knowledge would be useful to empower them to purposefully engage with data sharing and to protect themselves from undesirable practices [53-55].

Another theme that emerged in this study is related to *trade-offs* or *paying* with personal information for consumer mHealth app services. Participants not only agreed to trade personal information to receive good quality apps but also perceived commercial companies as providers of better quality apps than nonprofit and government agencies. At the same time, participants expressed helplessness related to the overwhelming nature of the digital sphere. It was normal to share personal information; it was part of the culture. Few interviewees agreed to use less visually appealing and functionally convenient mHealth apps if they were supported by nonprofit organizations and government agencies. This finding suggests that smartphone users rank usability, functionality, and visual appeal much higher than the credibility of the health information provided, despite the potential sensitivity of such information. This finding is particularly important in understanding the coping strategies employed by mobile app users. Therefore, it is important to explore this topic in future studies.

The match between an organization and an app cause or topic mattered. Participants were more willing to share personal information with mHealth apps if they perceived supporting organizations to be relevant to a health issue. The match between a supporting organization and an app was found on the abstract level of the organization's mission. Some participants expressed more enthusiasm downloading an mHealth app supported by a nonprofit organization, as it was believed to serve people's interests related to public health. Although the mission of nonprofit organizations was often viewed as helping people, government agencies were perceived as instruments that provide credible, research-based, and unbiased information. These findings contribute to our understanding of how the type of persuasion knowledge agent can influence consumers' perceptions of the agent's traits, competencies, and goals.

### Theoretical and Practical Implications

Some study findings echoed the results of previous studies. For example, we found that participants did not engage in reading app use agreements and privacy policies [10,12]. Expecting users to read large amounts of privacy information for a simple service such as an mHealth app is unrealistic, as it requires time and effort [10-12]. Previous research has shown that users often engage in trade-offs between being concerned about privacy of personal data and the potential benefits of using new technologies [10-13,19,20], whereas other studies have also suggested that users' decisions are simultaneously influenced by other factors, such as risk (privacy concern) and trust (perceived control) [56,57]. Similar to previous studies [12,16,35-37], this study showed that the majority of participants were less likely to trust a commercial organization than a nonprofit or governmental organization supporting an mHealth app.

To extend previous evidence, this study offers several novel findings. First, it explored the complexity of assigning self-promotion and public service motivations to the following different types of supporting organizations: commercial, governmental, and nonprofit. It applied the PKM theoretical framework to further our understanding of persuasion knowledge related to not only commercial companies but also noncommercial entities. The findings of this study will help government and nonprofit organizations to develop technology-based cues for smartphone users to make effortless, yet informed judgments about consumer mHealth app quality, credibility, and personal information protection. Such cues may involve an organization's logo and sector identification (for-profit [US \$], nonprofit [♥], and government [🏛️]). Cues can be created to identify the nature and the scope of personal and health information sharing. For example, such data can be labeled as those that are (1) not shared outside the app, (2) shared within the supporting organization, and (3) shared with third parties. The purpose of sharing could also be specified via the following visual cues: marketing and promotional message personalization, public health statistics, or other.

It might still be unclear whether solely user-centric approaches will be effective in educating users about information privacy protection, as shown by our findings that extend the scope of existing literature [58]. Thus, we suggest that through relevant

policy and grassroot efforts, leaders guarantee user autonomy and privacy when using consumer mHealth apps and increase the motivation of stakeholders to establish ethical rules of consumer mHealth app execution and personal data distribution for long-term success [59]. Initial efforts to voice concerns of mHealth users have recently been made by both nonprofit and government organizations. For instance, the Federal Trade Commission, in 2016, released guidelines about each app, requiring an explanation in plain language of the kind of data the app would collect, and who would have access to it [60]. In addition, a 2018 California law (with a compliance date of January 2020) that focuses on consumer apps that do not fall under the purview of the Health Insurance Portability and Accountability Act requires updated privacy policies and implementation of a consumer's right to erasure [61,62]. Furthermore, Xcertia, a nonprofit organization founded by the American Medical Association and other major health and technology organizations, released guidelines for mHealth apps that would help with the privacy and security of users' health information [63].

This study provides additional practical insights related to promotional mHealth apps. To build a positive brand-consumer relationship through new mobile communication channels, commercial companies should consider users' persuasion knowledge and advertising literacy levels because awareness of the supporting organization's true intentions, especially oriented toward serving the public with health management tools, may positively affect app and brand evaluations. From the findings of this study, we can conclude that commercial organizations specifically elicited mixed perceptions among participants. Commercially supported mHealth app quality was attractive to users; however, the primacy of self-promotional motives had negative connotations. It would be beneficial to continue the investigation of mixed attitudes toward commercial mHealth apps to identify situations when such apps are perceived as bad and misleading and when they are viewed in a positive light through users' personal experiences. Furthermore, it is necessary to think more carefully about a good match between an organization's focus and mission and an mHealth app topic. Otherwise, mHealth app promotional support could result in a negative perception of the supporting agency.

Although commercial support implies strong self-driven intentions and, thus, could elicit skepticism, government and nonprofit support could be associated with the goal of improving public health [37]. It is important to facilitate the support of mHealth apps by noncommercial organizations, as such organizations may leverage trust and perceptions of their digital products as highly credible, backed by research, and secure. Security, although not being the focus of the study, often emerged in participants' discussion of privacy, supporting previous literature that shows the intertwined (objective and perceived) nature of the 2 concepts [64-66]. Future studies will need to determine how different types of supporting organizations would affect users' perceptions of consumer mHealth apps at the same level of usability while also exploring the tipping point when the credibility of noncommercial apps

becomes more appealing to users than the usability of commercial apps.

Finally, it may be difficult to fully distinguish between the 2 types of mHealth apps: those that are designed to enhance public health and those that collect personal and basic health information for marketing purposes. Promotional apps do not exclude the purpose of providing high-quality health services. Instead, they might be characterized by purpose duality, where an organization's promotion happens by providing good mHealth services. The findings of this study indicate that purpose duality is assigned to mostly commercial organizations. However, it does not exclude the possibility that noncommercial entities may also use apps for promotion despite being more likely to be perceived as *having participants' best interests at heart*. Future studies should explore in detail the impact of purpose duality on users' perceptions and utilization of mHealth apps.

### Limitations

Although we gathered insights to understand smartphone users' persuasion knowledge of supporting organizations, it is not possible to generalize this study's findings to a larger population. Future research should use quantitative methods to make systematic comparisons with the standardization assumptions underlying probability statistics. Another limitation is that the findings could be dependent on interviewers' communication skills. As we used an in-depth interview method, interview questions themselves could be leading, which could affect participants' responses. The interview guide also included prompts for participants who did not offer much responses (eg, liking of the health app, perceptions of quality and credibility

of the health app, and decision to download or use the health app). Although we only used planned follow-up questions or probes that made interview questions more specific and helped direct the participants to the central issues of the study [67,68], it is necessary to conduct a replication study using the same methods but with a different sample to confirm the trustworthiness of the findings. In particular, probing might have led to participants' bias in thinking elaboratively about persuasion knowledge, which could have influenced the study's findings. In addition, social desirability could affect the participants' responses. Finally, given time and resource constraints, some thematic analysis procedures were not implemented but are highly advised for use in future qualitative work. These refer to soliciting feedback on final themes from participants and peer researchers not involved with the study and using other methods and secondary data to study the same phenomena.

### Conclusions

Smartphone users possess high levels of persuasion knowledge of supporting organizations and understand organizations' promotional intentions, especially those of commercial companies. However, in the complex mHealth app marketplace, users' cognitive capacity to scrutinize all relevant mHealth app cues is limited, which may result in undesirable consequences related to personal and health-sensitive information collection and management on the web by unauthorized parties. Although users understand the potential threats to their data privacy, they often *trade* their personal and health information for free, convenient, and *fun* mHealth app services. The discussion of and recommendations for the safe and ethical use and privacy of mHealth apps should continue.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

Interview questions.

[PDF File (Adobe PDF File), 101 KB - [mhealth\\_v9i4e16518\\_app1.pdf](#) ]

---

#### Multimedia Appendix 2

Table S1. Conceptual definitions of themes and examples of codes.

[PDF File (Adobe PDF File), 76 KB - [mhealth\\_v9i4e16518\\_app2.pdf](#) ]

---

#### Multimedia Appendix 3

Table S2. Participant demographics.

[PDF File (Adobe PDF File), 68 KB - [mhealth\\_v9i4e16518\\_app3.pdf](#) ]

---

### References

1. Martínez-Pérez B, de LTI, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. *J Med Internet Res* 2013;15(6):e120 [FREE Full text] [doi: [10.2196/jmir.2600](#)] [Medline: [23770578](#)]
2. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform* 2014;5(3):229 [FREE Full text] [doi: [10.5210/ojphi.v5i3.4814](#)] [Medline: [24683442](#)]
3. Chen J, Cade JE, Allman-Farinelli M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. *JMIR Mhealth Uhealth* 2015;3(4):e104 [FREE Full text] [doi: [10.2196/mhealth.4334](#)] [Medline: [26678569](#)]

4. Aitken M, Clancy B, Deanna N. The Growing Value of Digital Health - Evidence and Impact on Human Health and the Healthcare System. IQVIA. Parsippany, NJ: IQVIA Institute for Human Data Science; 2017 Nov. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health> [accessed 2019-10-05]
5. Lorenz T. Under Armour has acquired fitness apps MyFitnessPal and Endomondo for a combined \$560 million - Business Insider. Business Insider. 2015 Feb 04. URL: <http://www.businessinsider.com/under-armour-has-acquired-myfitnesspal-and-endomondo-for-a-combined-435-million-2015-2> [accessed 2016-03-31]
6. Clement J. Leading health and fitness apps in the U.S. 2018, by users Internet. Statista. 2019 Nov. URL: <https://www.statista.com/statistics/650748/health-fitness-app-usage-usa/> [accessed 2020-04-20]
7. Berthene A. BabyCenter grows its app strategy. Digital Commerce 360 Internet. 2016 May 16. URL: <https://www.digitalcommerce360.com/2016/05/16/babycenter-grows-its-app-strategy/> [accessed 2020-04-18]
8. Centers for Disease Control and Prevention. Mobile Apps | Mobile Activities | CDC. 2019. URL: <https://www.cdc.gov/mobile/mobileapp.html> [accessed 2019-08-19]
9. American Red Cross. Apps | American Red Cross. 2019. URL: <https://www.redcross.org/get-help/how-to-prepare-for-emergencies/mobile-apps.html> [accessed 2019-08-19]
10. Atienza AA, Zarcadoolas C, Vaughn W, Hughes P, Patel V, Chou WS, et al. Consumer Attitudes and Perceptions on mHealth Privacy and Security: Findings From a Mixed-Methods Study. *J Health Commun* 2015 Apr;20(6):673-679. [doi: [10.1080/10810730.2015.1018560](https://doi.org/10.1080/10810730.2015.1018560)] [Medline: [25868685](https://pubmed.ncbi.nlm.nih.gov/25868685/)]
11. Joeckel S, Dogruel L, Bowman N. The reliance on recognition and majority vote heuristics over privacy concerns when selecting smartphone apps among German and US consumers. *Information, Communication & Society* 2016 Jun 28;20(4):621-636 [FREE Full text] [doi: [10.1080/1369118X.2016.1202299](https://doi.org/10.1080/1369118X.2016.1202299)]
12. Kelley P, Cranor L, Sadeh N. Privacy As Part of the App Decision-making Process. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM; 2013 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27–May 2, 2013; Paris, France p. 3393-3402 URL: <https://doi.org/10.1145/2470654.2466466> [doi: [10.1145/2470654.2466466](https://doi.org/10.1145/2470654.2466466)]
13. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. *J Am Med Assoc* 2015 Apr;22(e1):e28-e33. [doi: [10.1136/amiainl-2013-002605](https://doi.org/10.1136/amiainl-2013-002605)] [Medline: [25147247](https://pubmed.ncbi.nlm.nih.gov/25147247/)]
14. Statista; 2019. Android & iOS free and paid apps share 2019 | Statista. URL: <https://www.statista.com/statistics/263797/number-of-applications-for-mobile-phones/> [accessed 2019-08-10]
15. Dehling T, Gao F, Schneider S, Sunyaev A. Exploring the Far Side of Mobile Health: Information Security and Privacy of Mobile Health Apps on iOS and Android. *JMIR Mhealth Uhealth* 2015 Jan 19;3(1):e8 [FREE Full text] [doi: [10.2196/mhealth.3672](https://doi.org/10.2196/mhealth.3672)] [Medline: [25599627](https://pubmed.ncbi.nlm.nih.gov/25599627/)]
16. John N, Peters B. Why privacy keeps dying: the trouble with talk about the end of privacy. *Information, Communication & Society* 2017 Feb 1;20(2):284-298 [FREE Full text] [doi: [10.1080/1369118x.2016.1167229](https://doi.org/10.1080/1369118x.2016.1167229)]
17. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. *J Med Syst* 2015 Jan;39(1):181 [FREE Full text] [doi: [10.1007/s10916-014-0181-3](https://doi.org/10.1007/s10916-014-0181-3)] [Medline: [25486895](https://pubmed.ncbi.nlm.nih.gov/25486895/)]
18. Meng W, Ding R, Chung S, Han S, Lee W. The Price of Free: Privacy Leakage in Personalized Mobile In-App Ads. In: *Proceedings of the Symposium on Network and Distributed System Security*. 2016 Presented at: Symposium on Network and Distributed System Security (NDSS); February 21-24, 2016; San Diego, CA. USA URL: [http://wenkewww.gtisc.gatech.edu/papers/ndss16\\_mobile\\_ad.pdf](http://wenkewww.gtisc.gatech.edu/papers/ndss16_mobile_ad.pdf) [doi: [10.14722/ndss.2016.23353](https://doi.org/10.14722/ndss.2016.23353)]
19. Park Y, Jang S. Understanding privacy knowledge and skill in mobile communication. *Computers in Human Behavior* 2014 Sep 01;38:296-303 [FREE Full text] [doi: [10.1016/j.chb.2014.05.041](https://doi.org/10.1016/j.chb.2014.05.041)]
20. Prince C. Do consumers want to control their personal data? Empirical evidence. *International Journal of Human-Computer Studies* 2018 Feb;110:21-32. [doi: [10.1016/j.ijhcs.2017.10.003](https://doi.org/10.1016/j.ijhcs.2017.10.003)]
21. Deitz G, Myers S, Stafford M. Understanding Consumer Response to Sponsorship Information: A Resource-Matching Approach. *Psychol Mark* 2012 Apr 01;29(4):226-239 [FREE Full text] [doi: [10.1002/mar.20517](https://doi.org/10.1002/mar.20517)]
22. Grand View Research, Inc. 2019 Jun. URL: <https://www.grandviewresearch.com/industry-analysis/mhealth-app-market> [accessed 2019-08-20]
23. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Geneva: World Health Organization; 2011. URL: <http://www.who.int/iris/handle/10665/44607> [accessed 2019-08-20]
24. Heart T, Ben-Assuli O, Shabtai I. A review of PHR. EMR and EHR integration: A more personalized healthcare and public health policy 2017 Mar 01;6(1):20-25 [FREE Full text] [doi: [10.1016/j.hlpt.2016.08.002](https://doi.org/10.1016/j.hlpt.2016.08.002)]
25. Privacy Rights Clearinghouse. Mobile Health and Fitness Apps: What Are the Privacy Risks? Internet. 2016 Dec 16. URL: <https://privacyrights.org/consumer-guides/mobile-health-and-fitness-apps-what-are-privacy-risks> [accessed 2020-08-20]
26. Krebs P, Duncan D. 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](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
27. Rozendaal E, Lapierre M, van RE, Buijzen M. Reconsidering Advertising Literacy as a Defense Against Advertising Effects. *Media Psychology* 2011 Oct;14(4):333-354 [FREE Full text] [doi: [10.1080/15213269.2011.620540](https://doi.org/10.1080/15213269.2011.620540)]
28. Friestad M, Wright P. The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *J CONSUM RES* 1994 Jun;21(1):1-31. [doi: [10.1086/209380](https://doi.org/10.1086/209380)]

29. Tutaj K, van Reijmersdal EA. Effects of online advertising format and persuasion knowledge on audience reactions. *Journal of Marketing Communications* 2012 Feb 1;18(1):5-18 [FREE Full text] [doi: [10.1080/13527266.2011.620765](https://doi.org/10.1080/13527266.2011.620765)]
30. Ahmed R, Beard F, Yoon D. Examining and Extending Advertising's Dual Mediation Hypothesis to a Branded Mobile Phone App. *Journal of Interactive Advertising* 2016 Jul 02;16(2):133-144 [FREE Full text] [doi: [10.1080/15252019.2016.1237315](https://doi.org/10.1080/15252019.2016.1237315)]
31. MacKenzie S, Lutz R. An Empirical Examination of the Structural Antecedents of Attitude toward the Ad in an Advertising Pretesting Context. *Journal of Marketing* 1989;53(2):48-65 [FREE Full text] [doi: [10.1177/002224298905300204](https://doi.org/10.1177/002224298905300204)]
32. Boerman S, van Reijmersdal EA, Neijens P. Sponsorship Disclosure effects of Duration on Persuasion Knowledge and Brand Responses. *J Commun* 2012 Dec 1;62(6):1047-1064 [FREE Full text] [doi: [10.1111/j.1460-2466.2012.01677.x](https://doi.org/10.1111/j.1460-2466.2012.01677.x)]
33. D'Astous A, Chartier F. A Study of Factors Affecting Consumer Evaluations and Memory of Product Placements in Movies. *Journal of Current Issues & Research in Advertising* 2000 Sep;22(2):31-40. [doi: [10.1080/10641734.2000.10505106](https://doi.org/10.1080/10641734.2000.10505106)]
34. de Pelsmacker P, Neijens PC. New advertising formats: How persuasion knowledge affects consumer responses. *Journal of Marketing Communications* 2012 Feb;18(1):1-4. [doi: [10.1080/13527266.2011.620762](https://doi.org/10.1080/13527266.2011.620762)]
35. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
36. Kim H. Pharmaceutical Companies as a Source of Health Information: A Pilot Study of the Effects of Source, Web Site Interactivity, and Involvement. *Health Marketing Quarterly* 2011 Feb 15;28(1):57-85 [FREE Full text] [doi: [10.1080/07359683.2011.545327](https://doi.org/10.1080/07359683.2011.545327)]
37. Szykman L, Bloom P, Blazing J. Does Corporate Sponsorship of a Socially-Oriented Message Make a Difference? An Investigation of the Effects of Sponsorship Identity on Responses to an Anti-Drinking and Driving Message. *Journal of Consumer Psychology* 2004 Jan 1;14(1):13-20 [FREE Full text] [doi: [10.1207/s15327663jcp1401&2\\_3](https://doi.org/10.1207/s15327663jcp1401&2_3)]
38. Dembosky A. Pregnancy apps raise fresh privacy concerns. *Financial Times Internet*. 2013 Sep 29. URL: <https://www.ft.com/content/1c560432-2782-11e3-ae16-00144feab7de> [accessed 2020-08-17]
39. Rosato D. What Your Period Tracker App Knows About You. *Consumer Reports*. 2020 Jan 28. URL: <https://www.consumerreports.org/health-privacy/what-your-period-tracker-app-knows-about-you/> [accessed 2020-08-17]
40. Steel E. US Senate expands data privacy investigation. *Financial Times*. 2013 Sep 25. URL: <https://www.ft.com/content/f4434964-25c4-11e3-8ef6-00144feab7de> [accessed 2020-08-17]
41. Westin AF. *Privacy and freedom*. New York, NY, USA: Atheneum; 1967:xvi-487.
42. Moskop J, Marco C, Larkin G, Geiderman J, Derse A. From Hippocrates to HIPAA: Privacy and confidentiality in Emergency Medicine? Part I: Conceptual, moral, and legal foundations. *Annals of Emergency Medicine* 2005 Jan;45(1):53-59 [FREE Full text] [doi: [10.1016/j.annemergmed.2004.08.008](https://doi.org/10.1016/j.annemergmed.2004.08.008)]
43. Becker M. Privacy in the digital age: comparing and contrasting individual versus social approaches towards privacy. *Ethics Inf Technol* 2019 Jul 17;21(4):307-317. [doi: [10.1007/s10676-019-09508-z](https://doi.org/10.1007/s10676-019-09508-z)]
44. Bellur S, Sundar S. How Can We Tell When a Heuristic Has Been Used? Design and Analysis Strategies for Capturing the Operation of Heuristics. *Communication Methods & Measures* 2014 Apr;8(2):116-137 [FREE Full text] [doi: [10.1080/19312458.2014.903390](https://doi.org/10.1080/19312458.2014.903390)]
45. Kanthawala S, Joo E, Kononova A, Peng W, Cotten S. Folk theorizing the quality and credibility of health apps. *Mobile Media & Communication* 2019 May 1;7(2):175-194 [FREE Full text] [doi: [10.1177/2050157918796859](https://doi.org/10.1177/2050157918796859)]
46. Chaiken S, Liberman A, Eagly A. Heuristic systematic information processing within beyond the persuasion context. In: Uleman JS, Bargh JA, editors. *Unintended thought*. New York, NY, USA: The Guilford Press; 1989:212-252.
47. Bowen G. Naturalistic inquiry and the saturation concept: a research note. In: *Qualitative research*. London, England, UK: Sage Publications; 2008:137-152.
48. Francis J, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles M, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology & Health* 2010 Dec;25(10):1229-1245 [FREE Full text] [doi: [10.1080/08870440903194015](https://doi.org/10.1080/08870440903194015)]
49. Boyatzis R. Transforming qualitative information. In: *Thematic analysis and code development*. USA: Sage Publications; Apr 1998.
50. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. *APA Handbook of Research Methods in Psychology* American Psychological Association. New York-London: Routledge; 2012:57-71.
51. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006 Apr;3(2):77-101 [FREE Full text] [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
52. Guest G, MacQueen K, Namey E. Validity and Reliability (Credibility and Dependability) in Qualitative Research and Data Analysis. In: *Applied Thematic Analysis*. Thousand Oaks, CA, USA: Sage Publications; 2012.
53. Leon P, Ur B, Wang Y, Sleeper M, Balebako R, Shay R, et al. What matters to users? Factors that affect users' willingness to share information with online advertisers. In: *Proceedings of the Ninth Symposium on Usable Privacy and Security*. 2013 Presented at: Symposium on Usable Privacy and Security; July 24–26, 2013; Newcastle, UK p. 1-12. [doi: [10.1145/2501604.2501611](https://doi.org/10.1145/2501604.2501611)]

54. Dehling T, Zhang Y, Sunyaev A. Consumer Perceptions of Online Behavioral Advertising. 2019 Presented at: IEEE 21st Conference on Business Informatics (CBI); July 15-17, 2019; Moscow, Russia p. 345-354. [doi: [10.1109/cbi.2019.00046](https://doi.org/10.1109/cbi.2019.00046)]
55. Varnali K. Online behavioral advertising: An integrative review. *Journal of Marketing Communications* 2019 Jun 17;27(1):93-114 [FREE Full text] [doi: [10.1080/13527266.2019.1630664](https://doi.org/10.1080/13527266.2019.1630664)]
56. Dinev T, Hart P. An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research* 2006 Mar;17(1):61-80 [FREE Full text] [doi: [10.1287/isre.1060.0080](https://doi.org/10.1287/isre.1060.0080)]
57. Buck C, Dinev T. Low Effort and Privacy ? How Textual Priming Affects Privacy Concerns of Email Service Users. In: *Proceedings of the 53rd Hawaii International Conference on System Sciences*. 2020 Presented at: Hawaii International Conference on System Sciences; January 1-20, 2020; Grand Wailea, Maui, USA URL: <https://hdl.handle.net/10125/64260> [doi: [10.24251/hicss.2020.518](https://doi.org/10.24251/hicss.2020.518)]
58. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. *Science American Association for the Advancement of Science* 2015 Jan 29;347(6221):509-514. [doi: [10.1126/science.aaa1465](https://doi.org/10.1126/science.aaa1465)]
59. Boerman S, Kruijkemeier S, Zuiderveen BF. Exploring Motivations for Online Privacy Protection Behavior: Insights From Panel Data. *Communication Research* 2018 Oct 05 [FREE Full text] [doi: [10.1177/0093650218800915](https://doi.org/10.1177/0093650218800915)]
60. FTC Releases New Guidance For Developers of Mobile Health Apps Internet. Federal Trade Commission. 2016. URL: <https://www.ftc.gov/news-events/press-releases/2016/04/ftc-releases-new-guidance-developers-mobile-health-apps> [accessed 2020-08-26]
61. Greene A. How a rushed California law will change the privacy and security landscape for mobile health apps. *MobiHealthNews*. 2018 Jul 27. URL: <https://www.mobihealthnews.com/content/how-rushed-california-law-will-change-privacy-and-security-landscape-mobile-health-apps> [accessed 2020-08-26]
62. Bill Text - AB-375 Privacy: personal information: businesses. California Legislative Information. URL: [https://leginfo.ca.gov/faces/billTextClient.xhtml?bill\\_id=201720180AB375](https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375) [accessed 2020-08-26]
63. Robeznieks A. Privacy, security guidelines aim to bolster use of mHealth Internet. American Medical Association. 2018. URL: <https://www.ama-assn.org/practice-management/digital/privacy-security-guidelines-aim-bolster-use-mhealth> [accessed 2020-08-26]
64. Balapour A, Nikkiah HR, Sabherwal R. Mobile application security: Role of perceived privacy as the predictor of security perceptions. *International Journal of Information Management* 2020 Jun 01;52:102063. [doi: [10.1016/j.ijinfomgt.2019.102063](https://doi.org/10.1016/j.ijinfomgt.2019.102063)]
65. Bonsón Ponte E, Carvajal-Trujillo E, Escobar-Rodríguez T. Influence of trust and perceived value on the intention to purchase travel online: Integrating the effects of assurance on trust antecedents. *Tourism Management* 2015 Apr 01;47:286-302. [doi: [10.1016/j.tourman.2014.10.009](https://doi.org/10.1016/j.tourman.2014.10.009)]
66. Rothstein MA. Health privacy in the electronic age. *J Leg Med* 2007 Dec 11;28(4):487-501 [FREE Full text] [doi: [10.1080/01947640701732148](https://doi.org/10.1080/01947640701732148)] [Medline: [18066975](https://pubmed.ncbi.nlm.nih.gov/18066975/)]
67. Bolderston A. Conducting a Research Interview. *Journal of Medical Imaging and Radiation Sciences* 2012 Mar 01;43(1):66-76 [FREE Full text] [doi: [10.1016/j.jmir.2011.12.002](https://doi.org/10.1016/j.jmir.2011.12.002)]
68. Burns K, McBride C, Patel B, FitzGerald G, Mathews S, Drennan J. Creating Consumer-Generated Health Data: Interviews and a Pilot Trial Exploring How and Why Patients Engage. *J Med Internet Res* 2019 Jun 13;21(6):e12367 [FREE Full text] [doi: [10.2196/12367](https://doi.org/10.2196/12367)]

## Abbreviations

- mHealth:** mobile health  
**PHI:** personal health information  
**PKM:** Persuasion Knowledge Model  
**RQ:** research question

*Edited by L Buis; submitted 06.10.19; peer-reviewed by E Willis, T Dehling; comments to author 25.11.19; revised version received 20.04.20; accepted 02.03.21; published 13.04.21.*

*Please cite as:*

Joo E, Kononova A, Kanthawala S, Peng W, Cotten S  
*Smartphone Users' Persuasion Knowledge in the Context of Consumer mHealth Apps: Qualitative Study*  
*JMIR Mhealth Uhealth* 2021;9(4):e16518  
URL: <https://mhealth.jmir.org/2021/4/e16518>  
doi: [10.2196/16518](https://doi.org/10.2196/16518)  
PMID: [33847596](https://pubmed.ncbi.nlm.nih.gov/33847596/)

©Eunsin Joo, Anastasia Kononova, Shaheen Kanthawala, Wei Peng, Shelia Cotten. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 13.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Acceptance, Barriers, and Future Preferences of Mobile Health Among Patients Receiving Trauma and Orthopedic Surgical Care: Paper-Based Survey in a Prospective Multicenter Study

Felix Reinecke<sup>1\*</sup>, Dr med; Florian Dittrich<sup>2,3\*</sup>, Dr med; Marcel Dudda<sup>1</sup>, Univ-Prof Dr med; Andreas Stang<sup>4</sup>, Univ-Prof Dr med; Christina Polan<sup>1</sup>, Dr med; Roman Müller<sup>1</sup>, Dr med; Paula Beck<sup>1</sup>, Dr med; Max Daniel Kauther<sup>1</sup>, Prof Dr med

<sup>1</sup>Department of Trauma, Hand and Reconstructive Surgery, University Medicine Essen, Essen, Germany

<sup>2</sup>Department of Orthopedics and Orthopedic Surgery, Saarland University Medical Center and Saarland University Faculty of Medicine, Homburg, Germany

<sup>3</sup>Joint Centre Bergischland, Sana Fabricius Clinic Remscheid, Remscheid, Germany

<sup>4</sup>Institute of Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany

\*these authors contributed equally

**Corresponding Author:**

Florian Dittrich, Dr med

Department of Orthopedics and Orthopedic Surgery

Saarland University Medical Center and Saarland University Faculty of Medicine

Kirrberger Straße, Building 37-38

Homburg, 66421

Germany

Phone: 49 68411624520

Email: [florian.dittrich@uks.eu](mailto:florian.dittrich@uks.eu)

## Abstract

**Background:** Smartphones have become an essential part of everyday life and it is undeniable that apps offer enormous opportunities for dealing with future challenges in public health. Nevertheless, the exact patient requirements for medical apps in the field of orthopedic and trauma surgery are currently unknown.

**Objective:** The aim of this study was to define target groups, evaluate patient requirements, and the potential and pitfalls regarding medical apps specific for patients receiving orthopedic and trauma surgical care.

**Methods:** A prospective multicenter study was conducted between August 2018 and December 2019 at a German trauma center and 3 trauma surgery/orthopedic practices. A paper-based survey consisting of 15 questions evaluated information regarding smartphone and medical app usage behavior. In addition, suggested app functions were rated using Likert scales. Descriptive statistics and binary log-binomial regression were performed.

**Results:** A total of 1055 questionnaires were included in our statistical analysis. Approximately 89.57% (945/1055) of the patients in this study owned a smartphone. Smartphone ownership probability decreased with every decade of life and increased with higher levels of education. Medical information was obtained via mobile web access by 62.65% (661/1055) of the patients; this correlated with smartphone ownership in regard to age and educational level. Only 11.18% (118/1055) of the patients reported previous medical app usage, and 3.50% (37/1055) of the patients received an app recommendation from a physician. More than half (594/1055, 56.30%) of the patients were unwilling to pay for a medical app. The highest rated app functions were *information about medication, behavioral guidelines, and medical record archival*. An improved treatment experience was reported through the suggested app features by 71.18% (751/1055) of the patients.

**Conclusions:** Mobile devices are a widely used source of information for medical content, but only a minority of the population reported previous medical app usage. The main target group for medical apps among patients receiving orthopedic and trauma surgical care tends to be the younger population, which results in a danger of excluding fringe groups, especially the older adults. Education seems to be one of the most important pull factors to use smartphones or a mobile web connection to obtain health information. Medical apps primarily focusing on an optimized patient education and flow of information seem to have the potential to support patients in health issues, at least in their subjective perception. For future target group-oriented app developments,

further evidence on the clinical application, feasibility, and acceptance of app usage are necessary in order to avoid patient endangerment and to limit socioeconomic costs.

(*JMIR Mhealth Uhealth* 2021;9(4):e23784) doi:[10.2196/23784](https://doi.org/10.2196/23784)

## KEYWORDS

smartphone; mHealth; app; orthopedics; trauma surgery

## Introduction

Today's health care professionals are faced with patients who are increasingly adapted to digitalization by using smartphones as tools for communication and information or data collection in their daily private and professional lives [1]. Before the era of the World Wide Web, patients often had only limited access to medical literature and therefore were completely dependent on the expertise of medical professionals. Ubiquitous access to the internet has fundamentally changed the information behavior regarding general knowledge of (non) medical issues for the majority of people [2]. This development has been enhanced by the widespread use of mobile web connections via smartphones. Smartphone ownership in Germany has increased steadily in recent years consecutively. In the first quarter of 2018, 87% of internet users had used smartphones or mobile devices to go online. Smartphone ownership and the associated mobile web usage has also risen globally [3,4].

Mobile health (mHealth) tools such as medical apps can enable patients to play a more active role in their health care [5]. In the past, health care services and medical information were often bound to medical facilities. Nowadays, by using mobile devices, a large target group can be reached to improve patient monitoring and self-engagement [4]. This offers the opportunity to address users who are otherwise difficult to access regarding health topics, such as older adults, younger people, or those living in rural regions with a low level of medical infrastructure [6,7]. In the course of the rapid development in the field of medical apps, not only have patients benefitted from this technology but physicians have also fundamentally changed their information behavior by using smartphones and apps to access web-based medical resources during their clinical routine in recent years [8]. The use of smartphones and medical apps seems to be very popular among trauma surgeons and orthopedic surgeons as well. In Germany, the majority (79.1%) of trauma and orthopedic surgeons reported the use of smartphones and medical apps (64.4%) in their daily clinical routine [9]. Despite the extensive possibilities arising from the use of this evolving technology, the evidence base is currently still limited and there is a need for further studies [10,11]. In orthopedic and trauma surgical care, apps can be used, for example, for preventing injuries. App-based training is able to prevent sport injuries such as anterior cruciate ligament sprains [12]. Patients with musculoskeletal pain such as chronic low back pain can also benefit from app support [13]. However, the integration and use of smartphones and medical apps in medical care, especially in the fields of orthopedic and trauma surgery, are still in an early developmental stage. Nevertheless, there seems to be numerous indications that mHealth solutions might have an additional benefit in the treatment of patients who have undergone

orthopedic and trauma surgery. The exact requirements and level of acceptance of medical apps from a patient's point of view are currently unknown. The target group for mHealth is also speculative and vaguely defined. Currently, the medical app features that are specifically important for these patients are still unknown. Therefore, it is essential to evaluate both the target group and patients' requirements for future patient-oriented medical app developments in the field of orthopedic and trauma surgery.

## Methods

### Study Design and Execution

This prospective, multicenter study was conducted between August 2018 and December 2019. Paper questionnaires were distributed to patients in a level 1 trauma center in western Germany (Essen University Hospital) and in 3 private practices with a focus on outpatient care in trauma and orthopedic surgery during their outpatient treatment in the facility. The questionnaires were handed out to the patients by the assistant staff at the patient registration desk. After giving their consent to participate in the study, patients were requested to drop the completed questionnaire in a labeled container. Participation was anonymous and on a facultative basis. All investigations on humans were carried out with the consent of the responsible ethics committee in accordance with national law and in accordance with the Declaration of Helsinki of 1975 (current revised version). Inclusion criteria were as follows: (1) patients undergoing ambulatory treatment in the aforementioned institutions, (2) patients aged  $\geq 15$  years and  $\leq 90$  years, and (3) existing consent for study participation. Exclusion criteria were as follows: (1) patients aged  $< 15$  years or  $> 90$  years and (2) missing declaration of consent for participation in the study

### Survey Development

As there is no gold standard for surveys in mHealth, a thorough literature review was conducted. App-related questions were developed and modified based on an already established survey [14]. The questionnaire was tested among a group of medical experts with know-how in the field of digitalism and survey development firstly and patients secondly. Based on feedback from the pretest survey, the final survey was created (Figure 1). A final questionnaire consisting of 15 questions was created. There were 3 sections in the survey. First, patients were asked about demographic characteristics (sex, age, educational background, insurance status, and type of treatment). The second part evaluated behavioral information regarding the patient's use of smartphones and medical apps. Additionally, patients were asked whether they had a smartphone, a mobile web connection, and if they used their internet access for medical research. In addition, it was evaluated if the study patients used

medical apps and whether they had ever received a recommendation for a medical app by a physician. Next, they were asked about their willingness to pay for an app in a medical context. Finally, patients were asked to rate the 10 proposed

features of a fictitious smartphone app on a 6-point Likert scale in order to determine their preferences for the app. Patients were also able to indicate whether they felt that they would benefit from these app features in terms of treatment experience.

**Figure 1.** Schematic illustration of the survey development process.



## Statistical Analysis

The completed questionnaires were returned to the investigator and given an identification number, which allowed conclusions about the institution collecting the data. The data were then entered into a Microsoft Excel worksheet (Version 15.18, Microsoft Corporation) and transferred to SAS 9.4 (Cary) for statistical analysis. Descriptive statistics were calculated for all items. In our statistical analysis, we classified patients with a university degree or a university (German) entrance qualification (13 years of school education) as having the highest level of education. Patients who had attended school for 10 years and stated that they had the corresponding German school certificate (“mittlere Reife”) were considered to have an average educational level. If the patients stated that they had no school diploma or the lowest German school certificate (9 years in school, “Hauptschulabschluss”), they were defined as having a low level of education. In addition, ranking of the best-rated app function was generated by ordering the mean ratings per app function. Ratings were on a scale of 1 (very important) to 6 (very unimportant). Rank 1 was considered to be the best-rated app with a consecutively low score. Since the app functions were evaluated using ordinal Likert scales, the median was chosen as the comparative location parameter. Furthermore, the interquartile range was determined. We performed Pearson chi-square test to determine whether there was a statistically significant association between the variables. All prevalence ratios and 95% confidence intervals were age-adjusted and sex-adjusted and were derived from log-binomial regression models.

## Results

### Pretest

Adjustments regarding content coherence, redundancy, and layout were made based on feedback from the pretest. Particular attention was paid to achieve only a short processing time for

the survey in order to keep the response rate as high as possible. The patients took an average of 80 seconds to complete the survey.

### Descriptive Statistics and Demographic Data

A total of 1331 questionnaires were distributed between August 2018 and December 2019. Of these, 1132 (85.05% response rate) were completed and returned. Seventy-seven patients did not meet the inclusion criteria. Therefore, 1055 questionnaires were included in the statistical evaluation. In the trauma center, 58.39% (616/1055) of the questionnaires were collected, while 41.61% (439/1055) of the questionnaires were obtained via the 3 private practices. The study consisted of 60.66% (640/1055) female patients and 39.34% (415/1055) male patients. The median age of the subjects was 45 years with an interquartile range of 30-59 years. Most patients were in the age group of 46-55 years (193/1055, 18.29%). Most of the patients had an average (314/1055, 29.76%) or high (511/1055, 48.44%) educational level; 19.24% (203/1055) of the patients reported a low educational status, while 2.56% (27/1055) of the patients were secondary school pupils. Most of the study patients (910/1055, 86.26%) had statutory health insurance, while 145 patients (13.74%) stated as having private health insurance. A total of 1015 (96.21%) patients were undergoing outpatient treatment, while 40 (3.79%) were being treated in an inpatient setting.

### Smartphone Usage Behavior

A large number of patients reported smartphone ownership (945/1055, 89.57%) and mobile web access (942/1055, 89.28%). A statistically significant correlation of smartphone ownership with higher educational level (odds ratio 1.13, 95% CI 5%-22%;  $P < .01$ ) and decreasing age (odds ratio  $-1.03$ , 95% CI 1.6%-3.2%;  $P < .01$ ) could be proved. More than half of the patients (661/1055, 62.65%) stated that they used their mobile web access to obtain medical information, while 394 (37.35%) patients did not use their mobile web access for this purpose. Almost half (164/334, 49.1%) of the older patients ( $\geq 56$  years)

reported that they searched for health information online compared to 69.3% (251/362) of the patients younger than 36 years. A significant link between using a mobile web connection for medical research and age was seen in the study cohort. With every increased decade of age, the probability of using mobile internet access for medical information decreased by 6% (95% CI 4%-8%,  $P < .01$ ) relatively. Furthermore, there was a correlation between higher level of education and the use of mobile web access for medical information. The probability of patients with average and high level of education to use mobile web access to obtain medical information was 40% (95% CI 16%-69%,  $P < .001$ ) and 70% (95% CI 43%-103%,  $P < .001$ ) higher, respectively, than of those with a low educational level. Only a few patients (118/1055, 11.18%) reported that they already use apps in a medical context, while 88.82% (937/1055) of the patients stated that they had not used medical apps yet. Again, the significant influence of age was obvious; with every 10-year age increase, the probability that patients used medical apps decreased by 16% (95% CI 8%-24%,  $P < .001$ ) relatively. In addition, this study revealed a connection between educational level and the use of medical apps. The probability for medical app usage among patients with an average or high educational

level was 92% (95% CI 0%-268%) or 116% (95% CI 16%-300%), respectively, greater than that among patients with a low educational level. Furthermore, we found a statistically significant association between medical app usage or web-based obtainment of medical information and the overall feeling of treatment improvement through the use of the app features we offered ( $P < .01$ ). Only a few patients had previously received app recommendations from a physician (37/1055, 3.51%), while 96.49% (1018/1055) of the patients had not received such recommendations (Table 1). When asked about the willingness to pay for a medical app, 56.30% (594/1055) of the patients were unwilling to pay money for medical apps, 10.71% (113/1055) of the patients were willing to pay up to €0.5 (US\$ 0.54), and 23.60% (249/1055) of the patients would spend up to €3 (US\$ 3.26). Almost 6.92% (73/1055) of the patients were willing to pay €7.50 (US\$ 8.16), followed by 2.46% (26/1055) of the patients who would pay up to €15 (US\$ 16.31) for a health app. The median was €1 (US\$ 1.08) with an interquartile range of €1-3 (US\$ 1.08-3.25). We identified a statistically significant correlation between higher patient age and an increased willingness to pay ( $P < .01$ ).

**Table 1.** Key data on smartphone and medical app usage behavior of the patients in this study (N=1055).

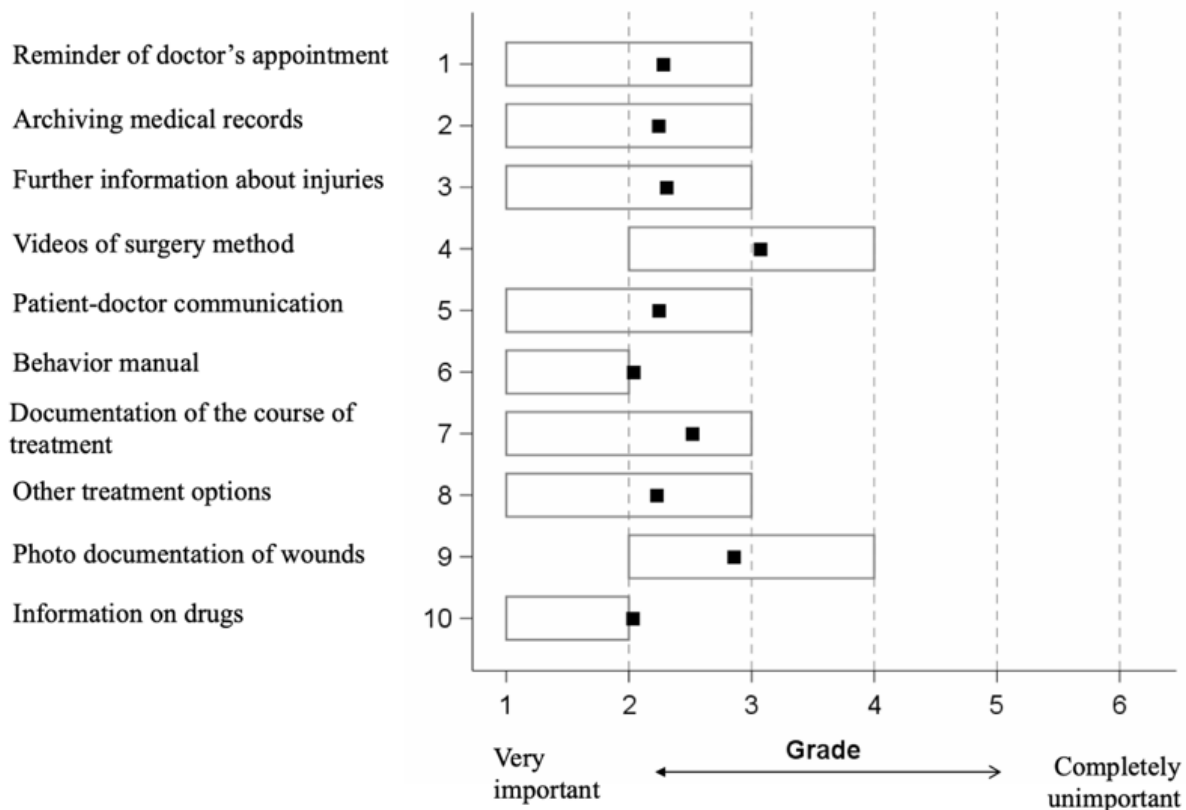
Usage behavior	Values, n (%)
<b>Smartphone ownership</b>	
Yes	945 (89.57)
No	110 (10.43)
<b>Mobile web access</b>	
Yes	942 (89.28)
No	113 (10.71)
<b>Is web access used to obtain medical information?</b>	
Yes	661 (62.65)
No	394 (37.35)
<b>Medical app use</b>	
Yes	118 (11.18)
No	937 (88.81)
<b>Has a medical app ever been recommended by a doctor?</b>	
Yes	37 (3.51)
No	1018 (96.49)

## App Functions

An app feature that provides a patient with information about the prescribed medication was rated best by patients in this survey (rank 1, score 1.91). An app providing behavior guidelines or discharge instructions following surgical procedures (eg, for traumas) was the second best-rated feature (rank 2, score 2.04). Archival of medical records was the third best-rated app function in this study (rank 3, score 2.19) (Figure 2). Across all age groups, 71.18% (751/1055) of the patients indicated that the aforementioned app features would improve

their treatment experience, while 28.82% (304/1055) negated this. Age had a significant influence on a perceived treatment improvement as a result of these app features because the app's functions were more likely to give younger patients a more positive feeling about their treatment. With every 10-year increase of age, the probability that patients would benefit from these app features decreased relatively by 4% (95% CI 2%-6%,  $P < .01$ ). However, even the majority of the patients older than 65 years (110/163, 67.5%) stated that they felt more comfortable in their treatment experience as a result of the app's features.

**Figure 2.** Evaluation of the features of a fictitious app by the patients in trauma surgical care (N=1055). The evaluation was based on a 6-point Likert scale from 1 (function very important) to 6 (function very unimportant). Boxes indicate interquartile ranges, and squares indicate median grades.



## Discussion

### Principal Findings

This study proved that smartphone usage and mobile web access are already widespread among patients receiving orthopedic and trauma surgical care. Mobile devices seem to be a widely used source of information for medical content. Education is one of the most important pull factors to use smartphones or a mobile web connection to obtain health information, which is in line with that reported in previous research [15-17]. Only a minority of the population stated the previous use of medical apps. Although many physicians use apps in their daily clinical routine, seemingly, not many consider medical apps to be important tools for patient care [8,9,18]. Consequently, very few patients reported having received a physician's recommendation for a medical app. Patients who reported that they had already used medical apps tended to be younger and had a higher educational level, which concurs with the findings of prior investigations [16,19]. As more than half of the patients were unwilling to pay for a medical app, a high sensitivity to the price of apps in a medical context might be apparent, which is consistent with that reported in previous studies [20,21]. More than 70% of the patients, including older adults, felt an additional benefit from the features suggested in the fictitious app presented. It is not surprising that younger people have a slightly greater app affinity as app users tend to be younger in general. Target group-specific apps, therefore, seem to have

the potential to support patients in health issues, at least in their subjective perception [22]. The most demanded app functions focused primarily on an optimized patient education and flow of information. *Information on drugs* was the highest rated app specification and might play an important role in pain management and self-management and therefore minimize drug-related complications [23,24]. A large number of medication management apps already exist, but none that is explicitly targeted at patients in a surgical specialty, where pain killers, especially opioids, are frequently prescribed [25-27]. *Posttraumatic and postoperative behavioral guidelines or discharge instructions* were also highly important (second highest rated). Former studies have already shown the prevalence of noncomprehensive discharge instructions, particularly among older adults who represent the major proportion of the trauma surgery patient clientele [28,29]. Important behavioral guidelines or other information may not be conveyed to the patient in a comprehensible manner [30]. In these instances, apps with implemented behavioral guidelines might support the patient in a postinterventional setting in addition to the conventional treatment [31].

### Limitations

This study has some limitations. This study was conducted solely in the field of orthopedics and trauma surgery in the German health care system. Because of this limited scope of application, the gained evidence is only valid for the respective target group. This necessitates the need to explore any additional

requirements that may be needed for medical apps in emerging nations or more rural nations and health care systems. Moreover, it is crucial to evaluate whether the requirements of patients for mHealth apps vary in disciplines other than orthopedic and trauma surgery. The results of this study are in line with the patient's requirements for apps in the field of chronic diseases of the musculoskeletal system. Using medical apps in rheumatology seems to be beneficial for the patient's outcome. However, the usage of mHealth among patients with rheumatism is very limited and eHealth literacy is rather poor too [32]. A paper-pencil-based survey was conducted to address all age groups equally, as older patients may not respond to a web-based survey. Future studies may choose a web-based survey to expand the number of patients. In addition, a new nonvalidated questionnaire was created for this study, which has not yet been used in large clinical trials. However, due to the pretest, disadvantages could be omitted. Since the study patients rated only 10 different suggested app features, it remains unclear whether other features might be of higher importance. Prospective app development must evaluate the importance of additional features.

### Outlook

Medical apps are a milestone in patient care and doctor-patient interaction. These apps are rapidly moving into health care and through further developments, might offer a wide range of possibilities for patients undergoing trauma surgery in the future. The patient structure in the field of orthopedic and trauma surgery is heterogeneous and is partly dominated by the older adults. One great challenge in the development of medical apps—also from an ethical point of view—is to find ways to address all patients equally. Future app developments should take target group-specific app requirements into serious consideration. As the best rated app features result in an optimized information flow, a remarkable information deficit or knowledge deficit during and after medical treatment might be inferred. Apps or other mHealth-based solutions may offer the opportunity to compensate for such deficiency. Information can literally be made available at the patients' fingertips and would be accessible at any time. It is questionable whether the majority of the patients reported a possible benefit from an app-supported treatment, as only a very small minority had ever received an mHealth app recommendation from a physician. After the initial ground-breaking steps, the German legislature recently gained considerable momentum in the direction of a

stringent national digitization strategy. The “Law for better care through digitization and innovation” (Digitale-Versorgung-Gesetz) passed by the Bundestag on November 7, 2019 paved the way for the prescription of apps, the improved use of web-based video consultation services, and greater data security in the communication of health data. This highlights the necessity to broaden the acceptance of this technology among treating physicians [33,34]. Given these possibilities, it will be necessary to gain well-founded evidence for the effectiveness of this technology in order to prevent high socioeconomic costs for inadequate apps and those that may endanger patients in Germany. Requirements such as data security and interoperability with clinical information systems will be mandatory to establish this technology in clinical routine and to increase its acceptance. As only a few patients in this study reported previous medical app usage, satisfactory or affordable offers seem to be lacking. Surgeons strive to support their patients by finding appropriate apps that address their specific needs. High-quality apps must be identified by involving physicians from corresponding specialties and professional associations [35]. Reimbursement programs might be helpful to broaden the use of apps, as our findings demonstrated a high price sensitivity. Medical professionals will be responsible to ensure that medical apps are not solely economically driven and have the primarily goal of improving patient health care. Future app developments should be based on medical guidelines and be accompanied by the expertise of medical professionals in order to create more transparency and benefit for patients.

### Conclusions

Mobile devices are a widely used source of information for medical content, but only a minority of the population reported previous medical app usage. The main target group for medical apps in orthopedic and trauma surgery tends to be the younger population, which results in a danger of excluding fringe groups, especially the older adults. Education seems to be one of the most important pull factors to use smartphones or a mobile web connection to obtain health information. Therefore, medical apps primarily focusing on an optimized patient education and flow of information seem to have the potential to support patients in health issues, at least in their subjective perception. For future target group-oriented app developments, further evidences on the clinical application, feasibility, and acceptance of app usage are necessary in order to avoid patient endangerment and limit socioeconomic costs.

---

### Acknowledgments

We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and Saarland University within the funding program Open Access Publishing.

---

### Conflicts of Interest

FD is active in Mediploy GmbH and in the development process of health apps. All other authors have no conflicts to declare.

---

### References

1. Dora J, van Hooff MLM, Geurts SAE, Hooftman WE, Kompier MAJ. Characterizing Work-Related Smartphone Use at Home and Private Smartphone Use at Work Using Latent Class Analysis. *Occup Health Sci* 2019 May 24;3(2):187-203. [doi: [10.1007/s41542-019-00040-6](https://doi.org/10.1007/s41542-019-00040-6)]

2. Bianco A, Zucco R, Nobile CGA, Pileggi C, Pavia M. Parents seeking health-related information on the Internet: cross-sectional study. *J Med Internet Res* 2013 Oct 18;15(9):e204 [FREE Full text] [doi: [10.2196/jmir.2752](https://doi.org/10.2196/jmir.2752)] [Medline: [24047937](https://pubmed.ncbi.nlm.nih.gov/24047937/)]
3. WeAreSocial. Digital in 2018: World's internet users pass the 4 billion mark. 2018. URL: <https://wearesocial.com/us/blog/2018/01/global-digital-report-2018> [accessed 2020-01-03]
4. Statistisches Bundesamt. Private Haushalte in der Informationsgesellschaft- Nutzung von Informations- und Kommunikationstechnologien. Fachserie 15 Reihe 4 zur IKT-Erhebung. 2020 Jan 16. URL: [https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publicationen/Downloads-IT-Nutzung/private-haushalte-ikt-2150400197004.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publicationen/Downloads-IT-Nutzung/private-haushalte-ikt-2150400197004.pdf?__blob=publicationFile) [accessed 2021-04-06]
5. Terry NP. Mobile health: assessing the barriers. *Chest* 2015 May;147(5):1429-1434. [doi: [10.1378/chest.14-2459](https://doi.org/10.1378/chest.14-2459)] [Medline: [25940254](https://pubmed.ncbi.nlm.nih.gov/25940254/)]
6. Rossmann C, Krömer N. mHealth in der medizinischen Versorgung, Prävention und Gesundheitsförderung. In: *eHealth in Deutschland Anforderungen und Potenziale innovativer Versorgungsstrukturen*. Heidelberg: Springer; 2016:441-456.
7. Abroms L, Padmanabhan N. *Mobile Phones For Health Communication To Promote Behavior Change*. NY, USA: Routledge; Aug 03, 2012:147-166.
8. Patel RK, Sayers AE, Patrick NL, Hughes K, Armitage J, Hunter IA. A UK perspective on smartphone use amongst doctors within the surgical profession. *Ann Med Surg (Lond)* 2015 Jul;4(2):107-112 [FREE Full text] [doi: [10.1016/j.amsu.2015.03.004](https://doi.org/10.1016/j.amsu.2015.03.004)] [Medline: [25905017](https://pubmed.ncbi.nlm.nih.gov/25905017/)]
9. Dittrich F, Back D, Harren A, Landgraerber S, Reinecke F, Serong S, et al. Smartphone and App Usage in Orthopedics and Trauma Surgery: Survey Study of Physicians Regarding Acceptance, Risks, and Future Prospects in Germany. *JMIR Form Res* 2020 Dec 30;4(11):e14787 [FREE Full text] [doi: [10.2196/14787](https://doi.org/10.2196/14787)] [Medline: [33252340](https://pubmed.ncbi.nlm.nih.gov/33252340/)]
10. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, 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](https://doi.org/10.2196/mhealth.8873)] [Medline: [29343463](https://pubmed.ncbi.nlm.nih.gov/29343463/)]
11. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Med* 2013;10(2):e1001382 [FREE Full text] [doi: [10.1371/journal.pmed.1001382](https://doi.org/10.1371/journal.pmed.1001382)] [Medline: [23424286](https://pubmed.ncbi.nlm.nih.gov/23424286/)]
12. van Mechelen DM, van Mechelen W, Verhagen EALM. Sports injury prevention in your pocket?! Prevention apps assessed against the available scientific evidence: a review. *Br J Sports Med* 2014 Jul;48(11):878-882. [doi: [10.1136/bjsports-2012-092136](https://doi.org/10.1136/bjsports-2012-092136)] [Medline: [23511697](https://pubmed.ncbi.nlm.nih.gov/23511697/)]
13. Chhabra HS, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. *Eur Spine J* 2018 Nov;27(11):2862-2874. [doi: [10.1007/s00586-018-5788-5](https://doi.org/10.1007/s00586-018-5788-5)] [Medline: [30324496](https://pubmed.ncbi.nlm.nih.gov/30324496/)]
14. Albrecht U, Afshar K, Illiger K, Becker S, Hartz T, Breil B, et al. Expectancy, usage and acceptance by general practitioners and patients: exploratory results from a study in the German outpatient sector. *Digit Health* 2017;3:2055207617695135 [FREE Full text] [doi: [10.1177/2055207617695135](https://doi.org/10.1177/2055207617695135)] [Medline: [29942582](https://pubmed.ncbi.nlm.nih.gov/29942582/)]
15. 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 Jul 16;16(7):e172 [FREE Full text] [doi: [10.2196/jmir.3117](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](https://pubmed.ncbi.nlm.nih.gov/25048379/)]
16. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmeier A, et al. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. *J Med Internet Res* 2017 Apr 05;19(4):e101 [FREE Full text] [doi: [10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)] [Medline: [28381394](https://pubmed.ncbi.nlm.nih.gov/28381394/)]
17. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. *J Med Internet Res* 2009 Mar 20;11(1):e4 [FREE Full text] [doi: [10.2196/jmir.1035](https://doi.org/10.2196/jmir.1035)] [Medline: [19275980](https://pubmed.ncbi.nlm.nih.gov/19275980/)]
18. Schleder S, Dendl LM, Niessen C, Stroszczynski C, Schreyer AG. [Acceptance of medical apps and e-books among German radiologists]. *Radiologe* 2017 Oct;57(9):752-759. [doi: [10.1007/s00117-017-0279-y](https://doi.org/10.1007/s00117-017-0279-y)] [Medline: [28707151](https://pubmed.ncbi.nlm.nih.gov/28707151/)]
19. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. *JMIR Mhealth Uhealth* 2015 Dec 04;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
20. Zia JK, Le T, Munson S, Heitkemper MM, Demiris G. Download Alert: Understanding Gastroenterology Patients' Perspectives on Health-Related Smartphone Apps. *Clin Transl Gastroenterol* 2015 Jul 02;6:e96 [FREE Full text] [doi: [10.1038/ctg.2015.25](https://doi.org/10.1038/ctg.2015.25)] [Medline: [26133109](https://pubmed.ncbi.nlm.nih.gov/26133109/)]
21. Datillo JR, Gittings DJ, Sloan M, Hardaker WM, Deasey MJ, Sheth NP. "Is There An App For That?" Orthopaedic Patient Preferences For A Smartphone Application. *Appl Clin Inform* 2017 Aug 16;8(3):832-844 [FREE Full text] [doi: [10.4338/ACI-2017-04-RA-0058](https://doi.org/10.4338/ACI-2017-04-RA-0058)] [Medline: [28885637](https://pubmed.ncbi.nlm.nih.gov/28885637/)]
22. Lu C, Hu Y, Xie J, Fu Q, Leigh I, Governor S, et al. The Use of Mobile Health Applications to Improve Patient Experience: Cross-Sectional Study in Chinese Public Hospitals. *JMIR Mhealth Uhealth* 2018 May 23;6(5):e126. [doi: [10.2196/mhealth.9145](https://doi.org/10.2196/mhealth.9145)] [Medline: [29792290](https://pubmed.ncbi.nlm.nih.gov/29792290/)]
23. Lovell MR, Luckett T, Boyle FM, Phillips J, Agar M, Davidson PM. Patient education, coaching, and self-management for cancer pain. *J Clin Oncol* 2014 Jul 01;32(16):1712-1720. [doi: [10.1200/JCO.2013.52.4850](https://doi.org/10.1200/JCO.2013.52.4850)] [Medline: [24799486](https://pubmed.ncbi.nlm.nih.gov/24799486/)]

24. Dayer L, Heldenbrand S, Anderson P, Gubbins P, Martin B. Smartphone medication adherence apps: Potential benefits to patients and providers. *Journal of the American Pharmacists Association* 2013 Mar;53(2):172-181 [FREE Full text] [doi: [10.1331/japha.2013.12202](https://doi.org/10.1331/japha.2013.12202)]
25. Chaudhary MA, Schoenfeld AJ, Harlow AF, Ranjit A, Scully R, Chowdhury R, et al. Incidence and Predictors of Opioid Prescription at Discharge After Traumatic Injury. *JAMA Surg* 2017 Oct 01;152(10):930-936 [FREE Full text] [doi: [10.1001/jamasurg.2017.1685](https://doi.org/10.1001/jamasurg.2017.1685)] [Medline: [28636707](https://pubmed.ncbi.nlm.nih.gov/28636707/)]
26. Finney FT, Gossett TD, Hu HM, Waljee J, Brummett C, Walton DM, et al. Conventional Ankle Sprain Treatment is Associated with Alarming High Rates of Persistent Opioid Use. *Foot & Ankle Orthopaedics* 2019 May 02;4(2):2473011419S0000. [doi: [10.1177/2473011419s00003](https://doi.org/10.1177/2473011419s00003)]
27. 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](https://doi.org/10.2196/13608)] [Medline: [31512580](https://pubmed.ncbi.nlm.nih.gov/31512580/)]
28. Makaryus AN, Friedman EA. Patients' understanding of their treatment plans and diagnosis at discharge. *Mayo Clin Proc* 2005 Aug;80(8):991-994. [doi: [10.4065/80.8.991](https://doi.org/10.4065/80.8.991)] [Medline: [16092576](https://pubmed.ncbi.nlm.nih.gov/16092576/)]
29. Albrecht JS, Gruber-Baldini AL, Hirshon JM, Brown CH, Goldberg R, Rosenberg JH, et al. Hospital discharge instructions: comprehension and compliance among older adults. *J Gen Intern Med* 2014 Dec;29(11):1491-1498 [FREE Full text] [doi: [10.1007/s11606-014-2956-0](https://doi.org/10.1007/s11606-014-2956-0)] [Medline: [25015430](https://pubmed.ncbi.nlm.nih.gov/25015430/)]
30. Dittrich F, Beck S, Busch A, Dudda M, Harren A, Sander A, et al. [Is the discharge letter still relevant? : Chances and risks of "Medical apps" in orthopedics and traumatology]. *Orthopade* 2018 Oct;47(10):842-848. [doi: [10.1007/s00132-018-3598-4](https://doi.org/10.1007/s00132-018-3598-4)] [Medline: [30039468](https://pubmed.ncbi.nlm.nih.gov/30039468/)]
31. Sens B, von Lewinski G, Albrecht U. "TEP-App" für die Endoprothetik. *Deutsches Aerzteblatt*. 2015. URL: [https://www.researchgate.net/publication/275644647\\_TEP-App\\_fur\\_die\\_Endoprothetik](https://www.researchgate.net/publication/275644647_TEP-App_fur_die_Endoprothetik) [accessed 2021-04-05]
32. 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](https://doi.org/10.2196/19661)] [Medline: [32678796](https://pubmed.ncbi.nlm.nih.gov/32678796/)]
33. Doctors should be able to prescribe apps. Bundesgesundheitsministerium. URL: <https://www.bundesgesundheitsministerium.de/digitale-versorgung-gesetz.html> [accessed 2020-01-05]
34. Burk E. Medical apps to become part of standard health care: Germany's federal ministry of health presents draft act. *Gleiss Lutz*. 2019. URL: <https://www.gleisslutz.com/node/18460/pdf> [accessed 2020-01-31]
35. Dittrich F, Beck S, Harren AK, Reinecke F, Serong S, Jung J, et al. Correction: Analysis of Secure Apps for Daily Clinical Use by German Orthopedic Surgeons: Searching for the "Needle in a Haystack". *JMIR Mhealth Uhealth* 2020 Jul 24;8(6):e21600 [FREE Full text] [doi: [10.2196/21600](https://doi.org/10.2196/21600)] [Medline: [32579538](https://pubmed.ncbi.nlm.nih.gov/32579538/)]

## Abbreviations

**mHealth:** mobile health

*Edited by G Eysenbach; submitted 23.08.20; peer-reviewed by J Knitza, H Akram, J Li; comments to author 01.09.20; revised version received 23.01.21; accepted 30.01.21; published 21.04.21.*

*Please cite as:*

*Reinecke F, Dittrich F, Dudda M, Stang A, Polan C, Müller R, Beck P, Kauther MD*

*Acceptance, Barriers, and Future Preferences of Mobile Health Among Patients Receiving Trauma and Orthopedic Surgical Care: Paper-Based Survey in a Prospective Multicenter Study*

*JMIR Mhealth Uhealth* 2021;9(4):e23784

URL: <https://mhealth.jmir.org/2021/4/e23784>

doi: [10.2196/23784](https://doi.org/10.2196/23784)

PMID: [33881401](https://pubmed.ncbi.nlm.nih.gov/33881401/)

©Felix Reinecke, Florian Dittrich, Marcel Dudda, Andreas Stang, Christina Polan, Roman Müller, Paula Beck, Max Daniel Kauther. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 21.04.2021. 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.



Original Paper

# Voice Interface Technology Adoption by Patients With Heart Failure: Pilot Comparison Study

Lida Anna Apergi<sup>1</sup>, PhD; Margret V Bjarnadottir<sup>1</sup>, PhD; John S Baras<sup>2</sup>, PhD; Bruce L Golden<sup>1</sup>, PhD; Kelley M Anderson<sup>3,4</sup>, PhD, RN, FNP, CHFNP-K; Jiling Chou<sup>4</sup>, MA; Nawar Shara<sup>3,4</sup>, PhD

<sup>1</sup>Robert H. Smith School of Business, University of Maryland, College Park, MD, United States

<sup>2</sup>Institute for Systems Research, University of Maryland, College Park, MD, United States

<sup>3</sup>Georgetown University, Washington, DC, United States

<sup>4</sup>Medstar Health Research Institute, Hyattsville, MD, United States

**Corresponding Author:**

Lida Anna Apergi, PhD

Robert H. Smith School of Business

University of Maryland

7621 Mowatt Ln

College Park, MD, 20742

United States

Phone: 1 301 405 3374

Email: [lapergi@umd.edu](mailto:lapergi@umd.edu)

## Abstract

**Background:** Heart failure (HF) is associated with high mortality rates and high costs, and self-care is crucial in the management of the condition. Telehealth can promote patients' self-care while providing frequent feedback to their health care providers about the patient's compliance and symptoms. A number of technologies have been considered in the literature to facilitate telehealth in patients with HF. An important factor in the adoption of these technologies is their ease of use. Conversational agent technologies using a voice interface can be a good option because they use speech recognition to communicate with patients.

**Objective:** The aim of this paper is to study the engagement of patients with HF with voice interface technology. In particular, we investigate which patient characteristics are linked to increased technology use.

**Methods:** We used data from two separate HF patient groups that used different telehealth technologies over a 90-day period. Each group used a different type of voice interface; however, the scripts followed by the two technologies were identical. One technology was based on Amazon's Alexa (Alexa+), and in the other technology, patients used a tablet to interact with a visually animated and voice-enabled avatar (Avatar). Patient engagement was measured as the number of days on which the patients used the technology during the study period. We used multiple linear regression to model engagement with the technology based on patients' demographic and clinical characteristics and past technology use.

**Results:** In both populations, the patients were predominantly male and Black, had an average age of 55 years, and had HF for an average of 7 years. The only patient characteristic that was statistically different ( $P=.008$ ) between the two populations was the number of medications they took to manage HF, with a mean of 8.7 (SD 4.0) for Alexa+ and 5.8 (SD 3.4) for Avatar patients. The regression model on the combined population shows that older patients used the technology more frequently (an additional 1.19 days of use for each additional year of age;  $P=.004$ ). The number of medications to manage HF was negatively associated with use ( $-5.49$ ;  $P=.005$ ), and Black patients used the technology less frequently than other patients with similar characteristics ( $-15.96$ ;  $P=.08$ ).

**Conclusions:** Older patients' higher engagement with telehealth is consistent with findings from previous studies, confirming the acceptability of technology in this subset of patients with HF. However, we also found that a higher number of HF medications, which may be correlated with a higher disease burden, is negatively associated with telehealth use. Finally, the lower engagement of Black patients highlights the need for further study to identify the reasons behind this lower engagement, including the possible role of social determinants of health, and potentially create technologies that are better tailored for this population.

(*JMIR Mhealth Uhealth* 2021;9(4):e24646) doi:[10.2196/24646](https://doi.org/10.2196/24646)

## KEYWORDS

heart failure; telehealth; voice interface; conversational agent; artificial intelligence; wireless technology; social determinants of health; mobile phone

## Introduction

### Background

Heart failure (HF) is a condition in which a patient's heart is unable to pump enough blood and oxygen to the organs. HF has high prevalence, affecting over 26 million people worldwide, and is associated with high mortality and health care utilization [1]. In the United States, there are currently 6.2 million adults with HF. Owing to the aging population, the number of individuals with HF is expected to exceed 8 million by 2030 (corresponding to approximately 2.97% of the US adult population) [2]. The total medical and indirect costs associated with HF are estimated to reach US \$70 billion by 2030 [3]. Currently, in the United States, there are approximately 800,000 annual hospitalizations for the primary diagnosis of HF, and after each hospitalization, the 28-day and 1-year mortality rates are 10.4% and 29.5%, respectively [2]. Thus, it is critical to support patients with HF in managing their conditions once they are discharged from the hospital.

The long-term management of HF is closely associated with self-care. In addition to taking medications, patients are advised to reduce salt and fluid intake [4]; monitor their weight daily; stay active through appropriate physical activity; and evaluate potential signs and symptoms such as swollen ankles, weight increase, or shortness of breath [5].

Telehealth offers potential benefits for patients with HF because it allows their health care providers to collect daily patient feedback and, therefore, enables them to promptly intervene when necessary. Several telehealth approaches to HF have been examined. Structured telephone monitoring allows patients to answer a set of prerecorded questions regarding their symptoms through their telephone keypad [6]. Increased internet access has further enabled the development of numerous technologies [7-11], with some examples presented later. Patients can log on to a designated website to enter information about their daily symptoms, which allows nurses to monitor any changes [11]. Through an Xbox gaming platform, patients with HF can navigate through screens, answer multiple-choice questions regarding their symptoms, read further instructions about their self-care, and learn more about their condition [8]. Tablets connected to a weight scale and a blood pressure wrist monitor allow patients to send daily readings to their health care provider [10]. Similarly, a number of recent studies have investigated smartphone apps that can be used by patients with HF to submit daily symptoms, transmit vital readings, and receive feedback about their health [7,12-16].

A recent review of studies on telehealth adoption by patients with HF can be found in the study by Gorst et al [17]. Across the studies discussed, the main factors identified as negative influences on telehealth adoption include difficulties with using the required technology, not remembering to use the technology every day, and considering the telehealth procedure to be redundant or boring. Therefore, it can be concluded that a critical

component of any telehealth application is its ease of use and engagement. Conversational agent technology using a voice interface is a potential solution because it can ask patients questions through speech and understand their answers through speech recognition.

Conversational agents have been used in numerous health care settings, with the literature extensively focusing on mental health applications [18,19]. However, conversational agent technology using a voice interface has also been used to help support behavior change and promote a healthy lifestyle [20]. Furthermore, in a few studies, the technology has been recommended for patients with HF to collect information about symptoms and management of their conditions. In particular, the proposed designs for voice interface technology for patients with HF can be found in the studies by Ferguson et al [21] and Zhang et al [22]. However, these studies did not provide results for evaluating the implementation of the proposed technology. One study on a small cohort of patients with HF, investigating the satisfaction and engagement with conversational agent technology, found high user satisfaction and an average engagement of about 60% [23]. However, the technology in this study was a chatbot that did not have a voice interface. Furthermore, this study had a small number of participants (5 patients) and did not examine how the characteristics of the patients impacted their level of engagement.

### Objectives

In this paper, we investigate which patient characteristics are associated with patient engagement with the voice interface technologies. To the best of our knowledge, this is the first study to examine the factors that influence HF patients' adoption of voice interface technology.

## Methods

### Voice Interface Technologies

We studied a voice technology using two different user interfaces, Alexa+ and a visual Avatar (both introduced in detail later). Data for the two technologies were collected from two different studies. The study of the Alexa+ technology was funded by the National Institutes of Health (trial registration number: NCT03707275), and the study for the Avatar was investigator initiated and industry funded. The two studies followed the exact design and protocol, which enabled us to compare user engagement with the 2 different voice interfaces and, more generally, to compare the drivers of their use.

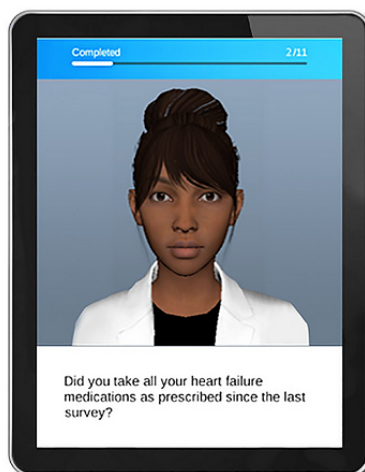
Alexa is a virtual assistant artificial intelligence technology developed by Amazon. Alexa is voice activated and has a number of functions, including sending messages, playing music, and providing traffic updates. For this study, the Alexa Skills Kit was used to expand the original capabilities of Alexa technology. A voice-activated survey was developed that enabled patients to answer a series of questions regarding their conditions and receive feedback. The resulting technology

(Alexa+) was implemented using Echo Dot devices, which are smart speakers that can be used to access Alexa. The retail price for an Echo Dot is US \$50 per unit; thus, this technology is relatively affordable.

The Avatar interface was developed by Oben [24]. The appearance of the Avatar was designed to be both reflective of

the care team from the patient's hospital and have characteristics of the patient population, where a majority of our patients were African Americans. The Avatar was programmed to ask the patients the same series of questions related to their HF treatment and symptoms and provide feedback. The Avatar app was saved on tablets that had no other apps. The avatar used in this study is shown in [Figure 1](#).

**Figure 1.** Avatar on a tablet.



Note that in addition to the way that information is presented, one difference between the two technologies is that the Alexa+ technology allows patients to set up daily reminders. Thus, patients participating in the Alexa+ study had the option to receive reminders at a particular time of day, which prompted them to answer the questionnaire. However, the tablets used in the Avatar study did not have this option.

Both Alexa+ and Avatar were populated with the same script. The detailed script, including the questions asked, their order, and the comments that the voice interfaces make in each case, can be found in [Multimedia Appendix 1](#). In the first stage of the script, the voice interface asked the patients with HF 11 questions, which can potentially be expanded to 13 questions, depending on the answers given by the patients. Patients answered each question separately, with yes or no. The answer that the patient provides to each question affects the type of comment or/and interjection that the voice interface makes right after the answer, as well as the next question. The 11-item questionnaire was divided into three components: compliance (questions 1-3), mild HF symptoms (questions 4-6), and moderate or severe HF symptoms (questions 7-11). The answers were captured, color coded, and displayed using a Tableau dashboard developed specifically for this study. The colors reflect the risk level of patients associated with each of the 3 components. The compliance questions evaluated whether a patient weighed themselves, took their HF medication, and avoided eating high-salt food. A negative answer to any of these questions raises a blue flag for that question. The mild HF symptom questions were designed to determine if the patient experienced shortness of breath with regular activity, had a cough, or had swollen ankles. If the patient gives a positive answer to any of the mild HF symptom questions, it raises an orange flag for the corresponding question. The moderate or severe HF symptom questions have to do with weight increase

and shortness of breath at rest or while sleeping. If the patient answered yes to any of the moderate or severe HF symptom questions, a red flag was raised. The raised flags can result in alerts, and any red flag affects the tailored response that the voice interface gives the patient once the questionnaire is completed. This is done in the second stage of the script, where the voice interface summarizes the answers provided by the patient and then provides the response.

The HF patient survey contained within the script was created based on the existing literature, including the literature on HF action plans, which were used for self-care management and symptom recognition [25-27], and the literature on previously developed telehealth systems that can help with the management of HF [28,29]. Color-coded alerts based on participants' answers were previously used in a telemanagement pilot study [28], where the system provided feedback based on a three-zone action plan. The three zones consisted of green (*stay in this zone*), yellow (*warning*), and red (*seek immediate help*). Similar color-coded zones were used in the studies by Vincent and Mutsch [30,31], where the proposed action plan consisted of green, yellow, and red zones. The green zone indicated that the patient was doing well, and patients were instructed to continue current medications, diet, and activities. The yellow zone indicated caution, and patients were advised to follow a low-sodium diet, take prescribed medications, and take an extra diuretic dose if necessary. The red zone indicated a medical emergency, and the patient was instructed to immediately obtain help.

### Patient Enrollment and Training

A total of 30 patients were enrolled in each of the two studies using the same eligibility criteria. All participants had at some point in the past either been admitted to the MedStar Washington Hospital Center for HF or had been seen in a MedStar Heart

Failure Clinic for HF. Furthermore, the patients were required to be aged 18 years or older and live in a house with Wi-Fi access. Finally, the patients could not participate in the studies if they had had a heart transplant or if they had a ventricular assist device. Participation was voluntary; patients who declined the offer to participate provided different reasons, including not wanting to take the daily surveys, not wanting another device in their lives, or lack of interest.

Participants did not receive any monetary incentives for participation; however, they were allowed to keep the Alexa and tablet devices after their participation in the study was completed. Participants were identified via electronic health records (EHRs) and by reviewing the schedules of providers at the aforementioned cardiology clinics. When potentially eligible participants were identified, the study coordinator introduced the study to them, discussed the details and logistics of the study and the risks and benefits, and allowed the participants to take time to make an informed decision about whether to participate. If they decided to take part, they signed the informed consent document, medical history information was collected, and the coordinator then proceeded with assigning them the corresponding technology.

Participants in the Alexa+ study were recruited from December 2018 to March 2019 and were provided a study-specific Amazon account, equipped with an Echo Dot configured to access the Alexa+ app. Participants were provided training on the Alexa Echo Dot, including Alexa voice training. The voice training consisted of a session of 25 phrase repetitions that allowed Alexa to improve its voice recognition capabilities for the target user. The patients participating in the Avatar study were recruited from February to December 2019 and were provided tablets with the Avatar app and were shown how to use the technology. In addition, during their training, patients participating in both studies completed their first questionnaire to ensure that the device was working properly and to answer any questions that the patients may have had.

### Demographic and Technology Survey

Before patients started using the technology that was assigned to them, they completed a demographic and technology survey. This allowed us to identify patient characteristics associated with low- or high-voice interface technology adoption in subsequent analyses. In particular, the demographic section of the survey included questions about age, gender, marital status, race, Hispanic heritage, annual household income, education, insurance coverage, number of years with HF, number of medications to manage HF, and visual impairment. The technology section of the survey included questions about the type of mobile phone, whether they used their phone to send

text messages, whether they accessed social media and browsed the internet on their phones, and how confident they felt using computers or other electronic devices.

### Monitoring

Once enrolled, patients were instructed to complete the questionnaire daily for 90 days. For participants who answered in a manner that indicated HF stability, the response was coded green and no alerts were generated. As discussed earlier, patients' responses could be deemed clinically undesirable if they raised any blue (compliance questions), orange (mild HF symptom questions), or red (moderate or severe HF symptom questions) flags. Red-flagged questions generated immediate text and email alerts to the study nurse, who monitored the alerts daily, including weekends and holidays. The texts contained the participant identification number and the following alert: "We have received a concerning daily response from patient [PatientID] that warrants your attention: [QuestionID] yes." In total, 281 alerts were generated for the Alexa+ group and 404 for the Avatar group. In each email alert, the answers were also summarized as a color display on the dashboard (Figure 2). Note that the gray color in the figure indicates skip logic, meaning these questions were not required based on the previous answers. Both questions have to do with medication compliance (Multimedia Appendix 1).

For each patient, a baseline was established based on their initial responses to the questionnaire. The study nurse reviewed alerts daily and evaluated each participant's stability compared with their baseline; any changes were evaluated based on each patient's initial answers.

Certain information gathered through our study was shared with the staff of the study institution. As part of the institution's general care practices, each patient with HF is assigned a nurse navigator, who is responsible for coordinating the clinical care of the participant. The nurse navigator at the study institution, the principal investigator, and the study physician were all informed of any change in the status of the patient. A change in status was defined as either a change from baseline in moderate or severe HF symptoms, multiple red flag responses, or persistent red flag responses.

Every 3 weeks, the study coordinator would contact participants who did not complete the questionnaire to check on the participant's status, reemphasize the importance of using the technology, and encourage completion. If needed, the study coordinator would provide additional training to the participants on how to initiate and communicate with the device. In some cases, participants would call and inform the study coordinator that they were traveling out of the state or country or on vacation and were unable to complete the questionnaire.

**Figure 2.** Example of color display used in the monitoring of patients, as it appears in the email sent to the study nurse. Green indicates that the patient either complied with the corresponding instruction (columns 1-5) or did not have the corresponding symptom (columns 6-13), orange indicates that the patient gave a positive answer to the corresponding mild heart failure (HF) symptom question, and red indicates that the patient gave a positive answer to the corresponding moderate or severe HF symptom question. Gray indicates skip logic, meaning these questions were not required based on the previous answers.



**Analysis**

We used descriptive statistics to characterize the demographic and clinical characteristics of the patients and their experience and confidence in using technology. In particular, we estimated the means and SD for the continuous variables and counts and corresponding percentages for the categorical variables. Furthermore, we report the number of missing values for each variable. We compared the patient populations participating in the Alexa+ study with that of the Avatar study based on the aforementioned characteristics. We used chi-square tests to compare the counts in the different categories between the two populations for each of the 12 categorical variables, and we used 2-sample *t* tests to test the differences between the two population means for the 3 continuous variables. We report the corresponding *P* values for each variable.

This analysis is based on data from demographic and technology surveys. However, in some cases, data regarding demographic and clinical characteristics were also retrieved from patients' EHRs to fill in missing values.

We defined the variable of interest in our study, patient engagement, as the number of days during the study period in which the patient used the corresponding voice interface to answer the questionnaire. We chose this variable because daily engagement with the technology is important for patients with HF, as their condition can deteriorate quickly. For example, weight gain overnight can be a sign of fluid retention. To capture such rapidly changing clinical signs and to promote successful self-care and health monitoring, patients were asked to use the technology daily. Our use of the number of days of compliance is similar to that in other studies in the literature [15,32]. We note that using the percentage or portion of days as the variable of interest would result in identical results and insights, as the study period is fixed. In addition, using the number of days facilitates easy interpretation of the regression model. We used visualizations to demonstrate how patient engagement changed over time in each study and to obtain an initial estimate of engagement levels.

We modeled the relationship between the engagement level and the patients' characteristics using multiple linear regression on the entire population (combining both studies). The dependent variable is the number of days that the patient used the voice interface technology during their own specific 90-day period. The independent variables include the demographic, clinical, and technology-related characteristics of the patient. In particular, we generated a binary variable for patients who had a smartphone, for patients who were very confident in using technology, for patients who had only a little or no confidence in using technology, for patients who had a college education or higher, for patients who had a high school education or less, for patients with an annual household income higher than US \$100,000, for patients with an annual household income less than US \$25,000, for patients who were married, for patients who were identified as Black, and for patients who were identified as White. Finally, to differentiate between the two studies, we generated a binary variable, taking the value of 1 for patients who were enrolled in the Avatar study and the value of 0 for patients enrolled in the Alexa+ study.

Given the small sample size, the variables were carefully selected. We chose to use the best subset approach to generate a model that explains the variation in the patient's engagement as well as possible while using as few variables as possible. In all, 15 control variables were measured at the time of enrollment. We noted that although the data were not normalized, all of the control variables were binary, with the exception of the number of medications used, years with HF, and age. We used the adjusted *R*<sup>2</sup> value for model selection. Like any reduced model, the interpretation of the coefficients may be biased if important variables are excluded from the model. A follow-up sensitivity analysis was conducted to help us understand the interaction between the type of technology and income level.

**Results**

**Population Statistics**

Of the 30 patients, 3 initially enrolled in the Avatar study were subsequently withdrawn because they could not be reached after

they provided their initial participation consent. [Table 1](#) shows the demographic, clinical, and technology-related characteristics of the patients participating in each study. As shown in the table, patients participating in the Alexa+ study were taking significantly more medications to manage their HF, with a mean of 8.7 (SD 4.0), as compared with patients in the Avatar study, with a mean of 5.8 (SD 3.4;  $P=.008$ ). For the remaining variables, there were no statistically significant differences between the two patient populations.

Overall, both populations were predominantly male (60% with Alexa+ and 63% with Avatar), Black (60% with Alexa+ and 63% with Avatar), had an average age of approximately 55 years (mean age 54 years, SD 11.7 for Alexa+ and 56.5 years, SD 12.1 for Avatar), and had had HF for an average of about 7 years (mean of 7.5, SD 8.1 for Alexa+ and 7.3, SD 6.4 for Avatar). Furthermore, the majority of patients in both studies had experience using smartphones and had confidence in using similar technology.

**Table 1.** Characteristics of patients participating in the two studies.

Characteristic	Alexa+ (n=30)	Avatar (n=27)	P value
<b>Age (years), mean (SD)</b>	54.0 (11.7)	56.5 (12.1)	.45
Missing, n (%)	2 (7)	1 (4)	
<b>Gender, n (%)</b>			>.99
Male	18 (60)	17 (63)	
Female	10 (33)	10 (37)	
Missing	2 (7)	0 (0)	
<b>Marital status, n (%)</b>			.80
Single, never married	7 (23)	6 (22)	
Married	11 (37)	15 (56)	
Living together, not married	3 (10)	1 (4)	
Separated or divorced or widowed	6 (20)	5 (19)	
Missing	3 (10)	0 (0)	
<b>Race, n (%)</b>			.54
Black	18 (60)	17 (63)	
Asian	0 (0)	1 (4)	
White	7 (23)	8 (30)	
Other	4 (13)	1 (4)	
Missing	1 (3)	0 (0)	
<b>Hispanic heritage, n (%)</b>			>.99
Yes	1 (3)	0 (0)	
No	27 (90)	27 (100)	
Missing	2 (7)	0 (0)	
<b>Annual household income (US \$), n (%)</b>			.17
0-25,000	10 (33)	5 (19)	
25,001-50,000	9 (30)	4 (15)	
50,001-100,000	2 (7)	5 (19)	
More than 100,000	5 (17)	8 (30)	
Missing	4 (13)	5 (19)	
<b>Education level, n (%)</b>			.49
Some high school or high school graduate	11 (37)	9 (33)	
Some college	10 (33)	7 (26)	
College graduate	3 (10)	2 (7)	
Postgraduate degree	2 (7)	6 (22)	
Missing	4 (13)	3 (11)	
<b>Years with HF<sup>a</sup>, mean (SD)</b>	7.5 (8.1)	7.3 (6.4)	.95
Missing, n (%)	5 (17)	8 (30)	
<b>Number of medications to manage HF, mean (SD)</b>	8.7 (4.0)	5.8 (3.4)	.008
Missing, n (%)	2 (7)	1 (4)	
<b>Visually impaired or blind, n (%)</b>			.11
Yes	4 (13)	0 (0)	
No	22 (73)	25 (93)	
Missing	4 (13)	2 (7)	

Characteristic	Alexa+ (n=30)	Avatar (n=27)	P value
<b>Type of mobile phone, n (%)</b>			.41
Basic	1 (3)	2 (7)	
Smart	25 (83)	22 (82)	
None	0 (0)	1 (4)	
Missing	4 (13)	2 (7)	
<b>Uses phone to text, n (%)</b>			>.99
Yes	25 (83)	24 (89)	
No	1 (3)	1 (4)	
Missing	4 (13)	2 (7)	
<b>Uses phone for social media, n (%)</b>			.58
Yes	15 (50)	17 (63)	
No	11 (37)	8 (30)	
Missing	4 (13)	2 (7)	
<b>Uses phone to browse the internet, n (%)</b>			.23
Yes	24 (80)	20 (74)	
No	2 (7)	5 (19)	
Missing	4 (13)	2 (7)	
<b>Confidence in using technology, n (%)</b>			.87
Very	11 (37)	12 (44)	
Somewhat	13 (43)	10 (37)	
Only a little	2 (7)	2 (7)	
Not at all	0 (0)	1 (4)	
Missing	4 (13)	2 (7)	

<sup>a</sup>HF: heart failure.

## Technology Engagement

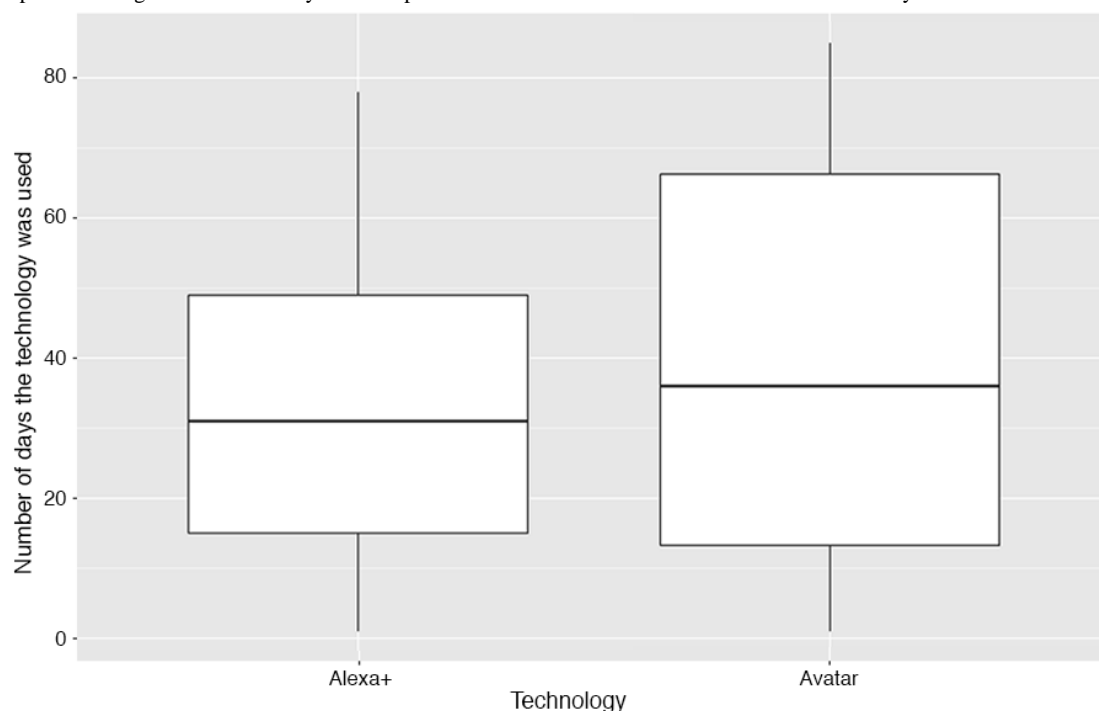
A few patients did not complete the study. In the Alexa+ study, 5 patients withdrew from the study, 2 patients had issues with their Wi-Fi connection, and 2 patients did not respond to contact attempts. Furthermore, 1 participant in the Avatar study did not respond to any contact attempts. Thus, from the Alexa+ study and the Avatar study, we obtained technology use information for 21 out of the 30 patients and 26 out of the 27 patients, respectively. The resulting study cohort is, therefore, 47 patients.

Alexa+ patients used the technology a mean of 35.3 times (SD 26.0), whereas the Avatar patients used it a mean of 37.8 times (SD 28.9). The *t* test of the difference between the two population means had a *P* value of .76, indicating that the difference in the engagement levels between the two groups

was not statistically significant. [Figure 3](#) highlights the variations in the level of engagement of patients participating in each study. For both technologies, we observed a large range of values in the number of times that patients interacted with the voice interface during the 90-day period. Although there were some patients in both studies who engaged with the technology almost daily, significantly less use of the technology was seen among the majority of participants.

Regarding the engagement over time for the two technologies, we observe a decrease in use over time. [Multimedia Appendices 2 and 3](#) show the use over time for the two technologies. The decrease in use was sharper for Alexa+ participants. On the other hand, participants in the Avatar study demonstrated more stable technology use in the second half of the study period.



**Figure 3.** Box plots showing the number of days that the patients interacted with the voice interface in each study.

### Regression Results

The regression model (Table 2) had an adjusted  $R^2$  of 28.13%. The resulting coefficients indicate that higher patient age is linked to higher use of the technology (1.19;  $P=.004$ ), whereas Black patients used the technology fewer times than non-Black patients with otherwise similar characteristics ( $-15.96$ ;  $P=.08$ ). Excluding other race indicator variables, the variable for the

Black race had a relatively strong correlation with the variable showing that the patient was married ( $-0.52$ ) and moderate correlation with an annual household income less than US \$50,000 (0.24) and years with HF ( $-0.22$ ). With the remaining variables, there were weaker correlations. Thus, this large difference in engagement of Black patients may be an indication of other important socioeconomic or medical factors or other confounding factors that are beyond the scope of our data.

**Table 2.** Linear regression model for predicting the number of times that the patient used the voice interface technology.

Variable	Coefficient	95% CI	<i>P</i> value
Intercept	17.93	-27.71 to 63.58	.43
Age	1.19	0.42 to 1.96	.004
Black	-15.96	-33.84 to 1.92	.08
Household income higher than US \$100,000	12.28	-5.49 to 30.05	.17
Confidence in using technology	12.88	-3.66 to 29.42	.12
Number of medications to manage HF <sup>a</sup>	-5.49	-9.22 to -1.72	.005
Avatar study participant	-24.14	-44.29 to -3.98	.02

<sup>a</sup>HF: heart failure.

The model also shows that patients who take a higher number of HF medications tended to demonstrate lower use of the technology ( $-5.49$ ;  $P=.005$ ). Finally, patients who participated in the Avatar study interacted less with the technology ( $-24.14$ ;  $P=.02$ ) compared with patients with similar characteristics who participated in the Alexa+ study. The direction of the Avatar regression coefficient is surprising when compared with Figure 3, which shows that the use of Avatar technology is higher on average. Therefore, we conducted a sensitivity analysis to understand what drives the coefficient value.

### Sensitivity Analysis

To further compare the participation levels in the Avatar and Alexa+ studies, we stratified participation by income levels. Figure 4 shows the engagement levels for each technology as a function of the income level. Note that the last group of boxplots corresponds to observations that have a missing value for household income. We see that although engagement with the technology increases as the household income increases in the case of the Alexa+ participants, the participation levels appear different for the Avatar participants. In particular, participants with a household income between US \$25,000 and US \$100,000 showed lower engagement than comparable

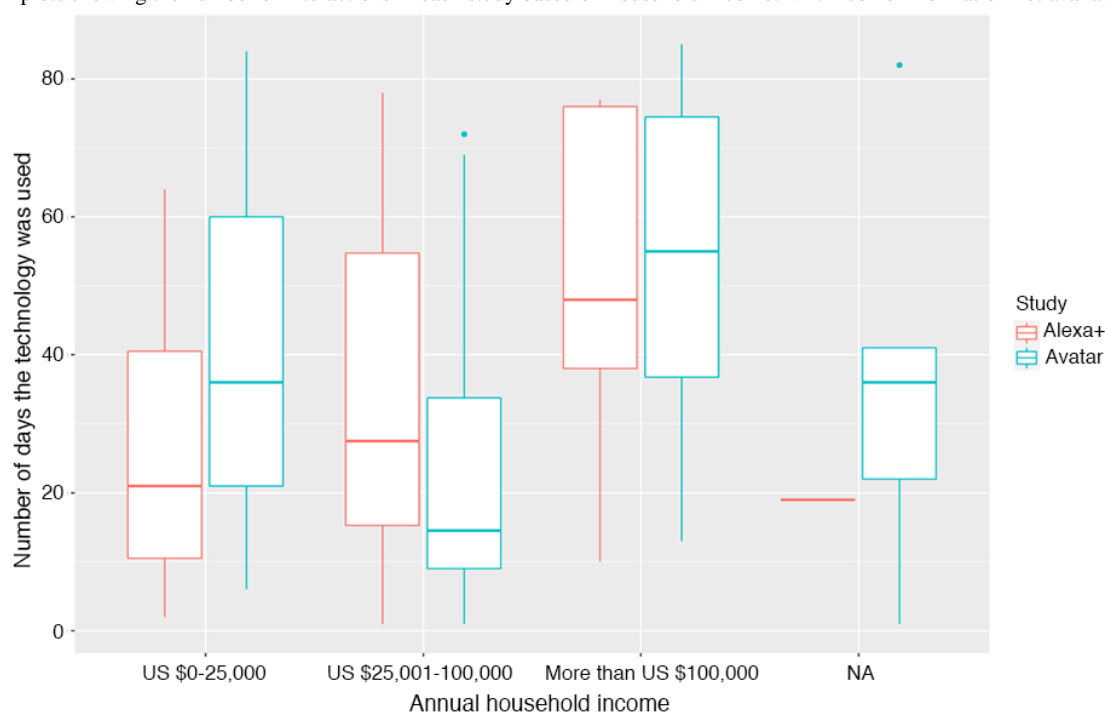
patients in the Alexa+ study. As income appears to have a different influence on engagement with the two technologies, we added interaction terms between Avatar and the different income levels and then reran the best subset regression.

The resulting regression model is shown in Table 3. The same overall patterns hold. Higher patient age was linked to higher use of the technology (1.07;  $P=.004$ ), whereas Black patients used the technology less frequently ( $-21.35$ ;  $P=.02$ ). Again, patients who take a higher number of medications to manage HF use the technology less on average ( $-6.52$ ;  $P=.002$ ).

From the model, we see that, overall, Avatar patients with middle- and high-income levels are expected to interact less with the technology compared with otherwise similar Alexa+ patients ( $-32.38$ ;  $P=.006$ ). However, the best subset regression

retained the interaction term between low income and Avatar, which, together with the Avatar coefficient, indicates that, on average, those from lower income households use Avatar technology more than otherwise similar Alexa+ participants. The interaction coefficient equals 44.81 ( $P=.04$ ), which, combined with the Avatar coefficient, estimates that Avatar participants from low-income households engage over 12 times more with the technology than similar Alexa+ participants. In other words, with everything else held constant, the model estimates that low-income patients use Avatar more often than Alexa+ technology, whereas the impact is reversed for middle-income patients and higher income patients. The overall impact of high and low household income was not statistically significant.

**Figure 4.** Box plots showing the number of interactions in each study based on household income. NA: income information not available.



**Table 3.** Linear regression model for predicting the number of times that the patient used the voice interface technology, with an added interaction term.

Variable	Coefficient	95% CI	P value
Intercept	43.45	2.54 to 84.36	.04
Age	1.07	0.37 to 1.78	.004
Black	-21.35	-39.19 to -3.50	.02
Household income higher than US \$100,000	15.19	-2.87 to 33.24	.10
Household income lower than US \$25,000	-3.89	-26.61 to 18.83	.73
Number of medications to manage HF <sup>a</sup>	-6.52	-10.33 to -2.70	.002
Avatar study participant	-32.38	-54.70 to -10.06	.006
Avatar×income lower than US \$25,000	44.81	2.44 to 87.18	.04

<sup>a</sup>HF: heart failure.

## Discussion

### Principal Findings

In this study, we compared the engagement levels of patients with HF with two different voice technology interfaces. To enable such a comparison between the two groups, the design and patient enrollment of both studies and the setup of both technologies were identical. Although the study was based on a small number of patients, the regression analysis helped identify some key characteristics that are linked to different engagement levels, contributing to the growing literature on technology use for chronic disease management [33].

On the basis of our analysis, Black patients used the technology 21 fewer times on average during the study period compared with non-Black patients with otherwise similar characteristics. As in our data set, Black race is not highly correlated with other features, this study adds to the evidence that technology design may need to be better tailored for this population [14]. Although our findings contribute to the emergent and growing literature on racial disparities in technology use, this area is still understudied. For instance, a recent survey on self-management among patients with HF points to only 4 studies that evaluate the use of technology for health management among Black patients with HF [34], the population that is arguably at the highest risk. Further examination to shed light on why Black patients engage less with health management technology is much needed, including the potential role of social determinants of health.

In addition, we see that patients taking a higher number of medications to manage HF interact less on average with the technology. The number of medications can be interpreted as a proxy for the severity of the HF condition or an indication of additional comorbid conditions. However, we note that the literature on the impact of disease burden on patient technology and self-care engagement is mixed; some studies have previously found that patients' willingness to self-monitor is not directly related to their health problems [35], whereas other studies indicate that sicker patients are more actively engaged [36] or less actively engaged [16,37]. The use of the number of medications for HF as a proxy for the severity of the condition has obvious limitations. A more direct measure of the severity of HF (eg, the New York Heart Association functional class [38] or the American College of Cardiology and American Heart Association stages of HF [39]) would be needed to more directly study the connection between patient engagement and disease severity in patients with HF.

Finally, our study finds that older patients tend to exhibit higher levels of engagement with voice interface technologies. This agrees with findings from previous studies that patient age affects the level of adherence, with older patients using telemonitoring more regularly [15,40-42].

### Future Research

The results presented here encourage future study of multiple aspects of voice interface technology, including its varying impact on and its potential to empower patients who are affected by different diseases. Specifically, important work can be done

by asking how best to operationalize these technologies and adapt them to different populations to support engagement. For example, one avenue for future work lies in advancing the technology's feedback: useful and personalized feedback is known to increase the level of engagement, and as a recent study points out, "The key to successful technology-based treatment is ease to use in a personalized manner such that ongoing feedback can be incorporated into these technology-based tools that keeps patients engaged" [43]. Although the technology used in this study did adjust its feedback in response to patient answers, more can be done to personalize such feedback. Future advances may draw on natural language technology, including voice interface technology, which is evolving quickly. Although this study used preset questions and feedback based on each patient's answers, the future of voice-assisted health monitoring may lie not in simple rule-based questions and answers but in open-ended questions to which a conversational agent can respond. In addition, it may be possible to better personalize voice interface technologies; for instance, drawing on the rich medical history data in EHR would allow us to create personalized voice interface technologies tailored to the specific patient's disease and preferences. Other potential avenues for increasing engagement include explaining the benefits of the technology and the details of its use. The engagement levels observed in this study highlight the importance of encouraging patients to use the technology by increasing awareness of its role in helping them manage their condition.

Our results further show the importance of future research into patient health factors associated with low levels of utilization and those that drive patient engagement. Further studies on larger sets of patients would allow for a more in-depth study of such patient engagement drivers, including the interaction between socioeconomic status and engagement levels. Such work could inform the development of voice interface systems, as this process should take patients' preferences, socioeconomic factors, and other population-level considerations into account. As a result, the potential benefits of voice interface apps can be more effectively leveraged for individuals and groups who are often underrepresented. In addition, larger studies of engagement drivers could allow for a better comparison between the drivers of engagement with voice interfaces and the drivers of engagement with other modes of self-monitoring. In general, a broader and deeper understanding of physiological and psychological factors that may prevent a patient from using a specific kind of technology would further the potential of that technology for patient self-management and early alerts of worsening condition.

This pilot study highlights the potential of patient engagement with a voice interface to help patients better manage their health; in addition to the larger issues of patient engagement discussed earlier, it opens up other future research possibilities. First, as this study compares two modes of voice technology, it does not shed light on the benefit of a voice interface over other self-monitoring technologies. In addition, this study was limited to patients with HF, whose daily self-monitoring is important (as, eg, change in weight overnight can be a sign of serious complication). Daily self-monitoring is also crucial for the management of many other chronic diseases, and the extent to

which our results transfer to other such chronic disease contexts is important and worthy of future study.

### Conclusions

HF is a chronic condition that requires sustainable management. Voice interface technologies present an opportunity to empower patients to better manage their health by setting reminders and following condition-specific instructions to potentially prevent

hospitalizations and emergency department visits. Our results show that an easy-to-use voice interface may play a significant role in helping older patients manage their chronic conditions and enhance remote engagement between patients and their providers to prevent worsening of the patient's condition. These innovative solutions have the potential to be scaled to address other chronic and more acute conditions and identify which subgroups of patients may best benefit from these technologies.

### Acknowledgments

This work was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number 3UL1TR001409-04S1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### Authors' Contributions

LA, MB, and NS contributed to the conception of the problem. KA and NS contributed to the design of the two voice interface studies and data acquisition. LA, MB, JC, and NS contributed to the data analysis and interpretation. LA drafted the manuscript. LA, MB, JB, BG, KA, and NS reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

Voice interface scripts.

[\[DOCX File , 105 KB - mhealth\\_v9i4e24646\\_app1.docx \]](#)

#### Multimedia Appendix 2

Percentage of Alexa+ participants (n=21) who logged in during each day in the study.

[\[PNG File , 62 KB - mhealth\\_v9i4e24646\\_app2.png \]](#)

#### Multimedia Appendix 3

Percentage of Avatar participants (n=26) who logged in during each day in the study.

[\[PNG File , 58 KB - mhealth\\_v9i4e24646\\_app3.png \]](#)

### References

1. Savarese G, Lund LH. Global Public Health Burden of Heart Failure. *Card Fail Rev* 2017 Apr;3(1):7-11 [[FREE Full text](#)] [doi: [10.15420/cfr.2016:25:2](https://doi.org/10.15420/cfr.2016:25:2)] [Medline: [28785469](https://pubmed.ncbi.nlm.nih.gov/28785469/)]
2. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. *Circulation* 2020 Mar 03;141(9):e139-e596. [doi: [10.1161/CIR.0000000000000757](https://doi.org/10.1161/CIR.0000000000000757)] [Medline: [31992061](https://pubmed.ncbi.nlm.nih.gov/31992061/)]
3. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail* 2013 May;6(3):606-619 [[FREE Full text](#)] [doi: [10.1161/HHF.0b013e318291329a](https://doi.org/10.1161/HHF.0b013e318291329a)] [Medline: [23616602](https://pubmed.ncbi.nlm.nih.gov/23616602/)]
4. Heart Disease, Heart Failure. Centers for Disease Control and Prevention. URL: <http://www.cdc.gov/heartdisease/heart> [accessed 2020-04-16]
5. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. *The Journal of cardiovascular nursing* 2009;24(6):485-497 [[FREE Full text](#)] [doi: [10.1097/JCN.0b013e3181b4baa0](https://doi.org/10.1097/JCN.0b013e3181b4baa0)] [Medline: [19786884](https://pubmed.ncbi.nlm.nih.gov/19786884/)]
6. Clark RA, Yallop JJ, Piterman L, Croucher J, Tonkin A, Stewart S, CHAT Study Team. Adherence, adaptation and acceptance of elderly chronic heart failure patients to receiving healthcare via telephone-monitoring. *European Journal of Heart Failure* 2007 Nov;9(11):1104-1111 [[FREE Full text](#)] [doi: [10.1016/j.ejheart.2007.07.018](https://doi.org/10.1016/j.ejheart.2007.07.018)] [Medline: [17942364](https://pubmed.ncbi.nlm.nih.gov/17942364/)]
7. Alluhaidan A, Lee E, Alnosayan N, Chatterjee S, Houston-Feenstra L, Dysinger W, et al. Designing patient-centered mHealth technology intervention to reduce hospital readmission for heart-failure patients. 2015 Jan 05 Presented at: 2015 48th Hawaii International Conference on System Sciences IEEE; 2015 Jan 5; Kauai, HI, USA p. 2886-2895. [doi: [10.1109/hicss.2015.349](https://doi.org/10.1109/hicss.2015.349)]

8. Finkelstein J, Wood J. Implementing home telemanagement of congestive heart failure using Xbox gaming platform. 2011 Aug 30 Presented at: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Boston, MA, USA p. 3158-3163. [doi: [10.1109/IEMBS.2011.6090861](https://doi.org/10.1109/IEMBS.2011.6090861)]
9. Kulshreshtha A, Kvedar JC, Goyal A, Halpern EF, Watson AJ. Use of remote monitoring to improve outcomes in patients with heart failure: a pilot trial. *International journal of telemedicine and applications* 2010;2010:870959 [FREE Full text] [doi: [10.1155/2010/870959](https://doi.org/10.1155/2010/870959)] [Medline: [20508741](https://pubmed.ncbi.nlm.nih.gov/20508741/)]
10. Lefler LL, Rhoads SJ, Harris M, Funderburg AE, Lubin SA, Martel ID, et al. Evaluating the Use of Mobile Health Technology in Older Adults With Heart Failure: Mixed-Methods Study. *JMIR Aging* 2018 Dec 04;1(2):e12178 [FREE Full text] [doi: [10.2196/12178](https://doi.org/10.2196/12178)] [Medline: [31518257](https://pubmed.ncbi.nlm.nih.gov/31518257/)]
11. Maric B, Kaan A, Araki Y, Ignaszewski A, Lear SA. The use of the Internet to remotely monitor patients with heart failure. *Telemedicine and e-Health* 2010;16(1):26-33. [doi: [10.1089/tmj.2009.0094](https://doi.org/10.1089/tmj.2009.0094)] [Medline: [20070163](https://pubmed.ncbi.nlm.nih.gov/20070163/)]
12. Alnosayan N, Chatterjee S, Alluhaidan A, Lee E, Houston FL. Design and Usability of a Heart Failure mHealth System: A Pilot Study. *JMIR Hum Factors* 2017 Mar 24;4(1):e9 [FREE Full text] [doi: [10.2196/humanfactors.6481](https://doi.org/10.2196/humanfactors.6481)] [Medline: [28341615](https://pubmed.ncbi.nlm.nih.gov/28341615/)]
13. Guo X, Gu X, Jiang J, Li H, Duan R, Zhang Y, et al. A Hospital-Community-Family-Based Telehealth Program for Patients With Chronic Heart Failure: Single-Arm, Prospective Feasibility Study. *JMIR Mhealth Uhealth* 2019 Dec 13;7(12):e13229 [FREE Full text] [doi: [10.2196/13229](https://doi.org/10.2196/13229)] [Medline: [31833835](https://pubmed.ncbi.nlm.nih.gov/31833835/)]
14. Heiney SP, Donevant SB, Arp Adams S, Parker PD, Chen H, Levkoff S. A Smartphone App for Self-Management of Heart Failure in Older African Americans: Feasibility and Usability Study. *JMIR Aging* 2020 Apr 03;3(1):e17142 [FREE Full text] [doi: [10.2196/17142](https://doi.org/10.2196/17142)] [Medline: [32242822](https://pubmed.ncbi.nlm.nih.gov/32242822/)]
15. Ware P, Dorai M, Ross HJ, Cafazzo JA, Laporte A, Boodoo C, et al. Patient Adherence to a Mobile Phone-Based Heart Failure Telemonitoring Program: A Longitudinal Mixed-Methods Study. *JMIR Mhealth Uhealth* 2019 Feb 26;7(2):e13259 [FREE Full text] [doi: [10.2196/13259](https://doi.org/10.2196/13259)] [Medline: [30806625](https://pubmed.ncbi.nlm.nih.gov/30806625/)]
16. Ware P, Ross HJ, Cafazzo JA, Boodoo C, Munnery M, Seto E. Outcomes of a Heart Failure Telemonitoring Program Implemented as the Standard of Care in an Outpatient Heart Function Clinic: Pretest-Posttest Pragmatic Study. *Journal of medical Internet research* 2020 Feb 08;22(2):e16538 [FREE Full text] [doi: [10.2196/16538](https://doi.org/10.2196/16538)] [Medline: [32027309](https://pubmed.ncbi.nlm.nih.gov/32027309/)]
17. Gorst SL, Armitage CJ, Brownsell S, Hawley MS. Home telehealth uptake and continued use among heart failure and chronic obstructive pulmonary disease patients: a systematic review. *Annals of Behavioral Medicine* 2014 Dec;48(3):323-336 [FREE Full text] [doi: [10.1007/s12160-014-9607-x](https://doi.org/10.1007/s12160-014-9607-x)] [Medline: [24763972](https://pubmed.ncbi.nlm.nih.gov/24763972/)]
18. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
19. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied Conversational Agents in Clinical Psychology: A Scoping Review. *Journal of medical Internet research* 2017 May 09;19(5):e151 [FREE Full text] [doi: [10.2196/jmir.6553](https://doi.org/10.2196/jmir.6553)] [Medline: [28487267](https://pubmed.ncbi.nlm.nih.gov/28487267/)]
20. Lisetti C, Yasavur U, De LC, Amini R, Visser U, Rishe N. Building an on-demand avatar-based health intervention for behavior change. 2012 May 25 Presented at: Flairs conference; 2012 May 23; Marco Island, FL, USA.
21. Ferguson G, Allen J, Galescu L, Quinn J, Swift M. Cardiac: An intelligent conversational assistant for chronic heart failure patient health monitoring. 2009 Presented at: 2009 AAAI Fall Symposium Series; 2009 Oct 30; Arlington, VA, USA.
22. Zhang L, Babu SV, Jindal M, Williams JE, Gimbel RW. A Patient-Centered Mobile Phone App (iHeartU) With a Virtual Human Assistant for Self-Management of Heart Failure: Protocol for a Usability Assessment Study. *JMIR research protocols* 2019 May 23;8(5):e13502 [FREE Full text] [doi: [10.2196/13502](https://doi.org/10.2196/13502)] [Medline: [31124472](https://pubmed.ncbi.nlm.nih.gov/31124472/)]
23. Moulik S. DIL-A Conversational Agent for Heart Failure Patients. The Claremont Graduate University. 2019. URL: <https://www.cgu.edu/event/dissertation-defense-sanjoy-moulik-phd-information-systems-and-technology/> [accessed 2020-07-02]
24. ObEN Artificial Intelligence. URL: <https://oben.me/> [accessed 2020-05-21]
25. Self-check plan for HF Management. Rise above heart failure. American Heart Association (AHA), Self-check plan for HF Management. 2015. URL: <https://www.heart.org/-/media/files/health-topics/heart-failure/self-check-plan-for-hf-management-477328.pdf?la=en&hash=875F6CDFB9A16372FA8CABC3552D14EABFEFABCA> [accessed 2020-08-05]
26. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. *European journal of heart failure* 2011 Feb;13(2):115-126. [doi: [10.1093/eurjhf/hfq219](https://doi.org/10.1093/eurjhf/hfq219)] [Medline: [21148593](https://pubmed.ncbi.nlm.nih.gov/21148593/)]
27. White MF, Kirschner J, Hamilton MA. Self-care guide for the heart failure patient. *Circulation* 2014 Jan 21;129(3):e293-e294. [doi: [10.1161/CIRCULATIONAHA.113.003991](https://doi.org/10.1161/CIRCULATIONAHA.113.003991)] [Medline: [24446413](https://pubmed.ncbi.nlm.nih.gov/24446413/)]
28. Finkelstein J, Dennison C. A pilot study of home automated telemanagement (HAT) system in African Americans with Congestive Heart Failure. 2010 Presented at: 2010 Second International Conference on eHealth, Telemedicine, and Social Medicine. IEEE; 2010 Feb 10; Netherlands Antilles. [doi: [10.1109/etelemed.2010.32](https://doi.org/10.1109/etelemed.2010.32)]
29. Finkelstein J, Wood J. Delivering chronic heart failure telemanagement via multiple interactive platforms. *management* 2013;6:8.

30. Vincent SE, Mutsch KS. Exploration of Self-Care Following Distribution of Acute Management Tool for Elder Heart Failure Patients in Clinic Setting. *Clinical Scholars Review* 2015 Jan 01;8(2):215-221. [doi: [10.1891/1939-2095.8.2.215](https://doi.org/10.1891/1939-2095.8.2.215)]
31. Vincent SE, Mutsch KS. Recognizing heart failure symptoms can improve patient self-management. *American Nurse Today* 2015;10(2):1-5.
32. Bickmore T, Schulman D, Yin L. Maintaining Engagement in Long-term Interventions with Relational Agents. *Applied Artificial Intelligence* 2010 Jul 01;24(6):648-666 [FREE Full text] [doi: [10.1080/08839514.2010.492259](https://doi.org/10.1080/08839514.2010.492259)] [Medline: [21318052](https://pubmed.ncbi.nlm.nih.gov/21318052/)]
33. Jiang J, Cameron A. IT-Enabled Self-Monitoring for Chronic Disease Self-Management: An Interdisciplinary Review. *MIS Quarterly* 2020 Jan 01;44(1):451-508. [doi: [10.25300/misq/2020/15108](https://doi.org/10.25300/misq/2020/15108)]
34. Hughes HA, Granger BB. Racial disparities and the use of technology for self-management in blacks with heart failure: a literature review. *Current heart failure reports* 2014 Sep;11(3):281-289 [FREE Full text] [doi: [10.1007/s11897-014-0213-9](https://doi.org/10.1007/s11897-014-0213-9)] [Medline: [25012939](https://pubmed.ncbi.nlm.nih.gov/25012939/)]
35. Huygens MWJ, Swinkels ICS, de Jong JD, Heijmans MJWM, Friele RD, van Schayck OCP, et al. Self-monitoring of health data by patients with a chronic disease: does disease controllability matter? *BMC family practice* 2017 Mar 20;18(1):40 [FREE Full text] [doi: [10.1186/s12875-017-0615-3](https://doi.org/10.1186/s12875-017-0615-3)] [Medline: [28320330](https://pubmed.ncbi.nlm.nih.gov/28320330/)]
36. Seng EK, Prieto P, Boucher G, Vives-Mestres M. Anxiety, Incentives, and Adherence to Self-Monitoring on a Mobile Health Platform: A Naturalistic Longitudinal Cohort Study in People With Headache. *Headache: The Journal of Head and Face Pain* 2018 Nov;58(10):1541-1555 [FREE Full text] [doi: [10.1111/head.13422](https://doi.org/10.1111/head.13422)] [Medline: [30334248](https://pubmed.ncbi.nlm.nih.gov/30334248/)]
37. Karhula T, Vuorinen A, Rääpysjärvi K, Pakanen M, Itkonen P, Tepponen M, et al. Telemonitoring and Mobile Phone-Based Health Coaching Among Finnish Diabetic and Heart Disease Patients: Randomized Controlled Trial. *Journal of medical Internet research* 2015;17(6):e153 [FREE Full text] [doi: [10.2196/jmir.4059](https://doi.org/10.2196/jmir.4059)] [Medline: [26084979](https://pubmed.ncbi.nlm.nih.gov/26084979/)]
38. Criteria Committee of the New York Heart Association. Nomenclature and criteria for diagnosis of diseases of the heart and great vessels. Boston: Little, Brown & Co; 1994.
39. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, et al. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2013 Oct 15;128(16):1810-1852 [FREE Full text] [doi: [10.1161/CIR.0b013e31829e8807](https://doi.org/10.1161/CIR.0b013e31829e8807)] [Medline: [23741057](https://pubmed.ncbi.nlm.nih.gov/23741057/)]
40. Krukowski RA, Harvey-Berino J, Bursac Z, Ashikaga T, West DS. Patterns of success: online self-monitoring in a web-based behavioral weight control program. *Health Psychology* 2013 Feb;32(2):164-170. [doi: [10.1037/a0028135](https://doi.org/10.1037/a0028135)] [Medline: [22545978](https://pubmed.ncbi.nlm.nih.gov/22545978/)]
41. Piette JD, Rosland A, Marinec NS, Striplin D, Bernstein SJ, Silveira MJ. Engagement with automated patient monitoring and self-management support calls: experience with a thousand chronically ill patients. *Medical Care* 2013 Mar;51(3):216-223 [FREE Full text] [doi: [10.1097/MLR.0b013e318277ebf8](https://doi.org/10.1097/MLR.0b013e318277ebf8)] [Medline: [23222527](https://pubmed.ncbi.nlm.nih.gov/23222527/)]
42. Steinberg DM, Levine EL, Lane I, Askew S, Foley PB, Puleo E, et al. Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among Black women: randomized controlled trial. *Journal of medical Internet research* 2014;16(4):e114 [FREE Full text] [doi: [10.2196/jmir.2996](https://doi.org/10.2196/jmir.2996)] [Medline: [24780934](https://pubmed.ncbi.nlm.nih.gov/24780934/)]
43. Dinesen B, Nonnecke B, Lindeman D, Toft E, Kidholm K, Jethwani K, et al. Personalized Telehealth in the Future: A Global Research Agenda. *Journal of medical Internet research* 2016 Mar 01;18(3):e53 [FREE Full text] [doi: [10.2196/jmir.5257](https://doi.org/10.2196/jmir.5257)] [Medline: [26932229](https://pubmed.ncbi.nlm.nih.gov/26932229/)]

## Abbreviations

**EHR:** electronic health record

**HF:** heart failure

*Edited by L Buis; submitted 28.09.20; peer-reviewed by H Tanaka, E Bellei; comments to author 09.11.20; revised version received 27.12.20; accepted 19.02.21; published 01.04.21.*

### *Please cite as:*

Apergi LA, Bjarnadottir MV, Baras JS, Golden BL, Anderson KM, Chou J, Shara N  
*Voice Interface Technology Adoption by Patients With Heart Failure: Pilot Comparison Study*  
*JMIR Mhealth Uhealth* 2021;9(4):e24646  
URL: <https://mhealth.jmir.org/2021/4/e24646>  
doi: [10.2196/24646](https://doi.org/10.2196/24646)  
PMID: [33792556](https://pubmed.ncbi.nlm.nih.gov/33792556/)

©Lida Anna Apergi, Margret V Bjarnadottir, John S Baras, Bruce L Golden, Kelley M Anderson, Jiling Chou, Nawar Shara. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 01.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

# Barriers to Use of Remote Monitoring Technologies Used to Support Patients With COVID-19: Rapid Review

Elizabeth Houlding<sup>1,2,3</sup>, HBSc; Kedar K V Mate<sup>3</sup>, PhD, MScPT; Kim Engler<sup>3</sup>, PhD; David Ortiz-Paredes<sup>3</sup>, MD, MSc; Marie-Pascale Pomey<sup>4,5</sup>, MD, PhD; Joseph Cox<sup>1,3,6</sup>, MD, MSc; Tarek Hijal<sup>7</sup>, MD, MSc, CM; Bertrand Lebouche<sup>1,3,8</sup>, MD, PhD

<sup>1</sup>Chronic Viral Illness Service, Royal Victoria Hospital, McGill University Health Centre, Montréal, QC, Canada

<sup>2</sup>Department of Physical Therapy, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

<sup>3</sup>Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montréal, QC, Canada

<sup>4</sup>Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montréal, QC, Canada

<sup>5</sup>Département de gestion, évaluation et politique de santé, École de santé publique de l'Université de Montréal, Montréal, QC, Canada

<sup>6</sup>Department of Epidemiology and Biostatistics, Faculty of Medicine, McGill University, Montréal, QC, Canada

<sup>7</sup>Division of Radiation Oncology, McGill University Health Centre, Montréal, QC, Canada

<sup>8</sup>Department of Family Medicine, McGill University, Montréal, QC, Canada

**Corresponding Author:**

Bertrand Lebouche, MD, PhD

Chronic Viral Illness Service

Royal Victoria Hospital

McGill University Health Centre

Local D02.4110, Glen Site

1001 Décarie Blvd

Montréal, QC, H4A 3J1

Canada

Phone: 1 514 398 7375

Email: [bertrand.lebouche@gmail.com](mailto:bertrand.lebouche@gmail.com)

## Abstract

**Background:** The COVID-19 pandemic has acted as a catalyst for the development and adoption of a broad range of remote monitoring technologies (RMTs) in health care delivery. It is important to demonstrate how these technologies were implemented during the early stages of this pandemic to identify their application and barriers to adoption, particularly among vulnerable populations.

**Objective:** The purpose of this knowledge synthesis was to present the range of RMTs used in delivering care to patients with COVID-19 and to identify perceived benefits of and barriers to their use. The review placed a special emphasis on health equity considerations.

**Methods:** A rapid review of published research was conducted using Embase, MEDLINE, and QxMD for records published from the inception of COVID-19 (December 2019) to July 6, 2020. Synthesis involved content analysis of reported benefits of and barriers to the use of RMTs when delivering health care to patients with COVID-19, in addition to health equity considerations.

**Results:** Of 491 records identified, 48 publications that described 35 distinct RMTs were included in this review. RMTs included use of existing technologies (eg, videoconferencing) and development of new ones that have COVID-19-specific applications. Content analysis of perceived benefits generated 34 distinct codes describing advantages of RMTs, mapped to 10 themes overall. Further, 52 distinct codes describing barriers to use of RMTs were mapped to 18 themes. Prominent themes associated with perceived benefits included a lower burden of care (eg, for hospitals, health care practitioners; 28 records), reduced infection risk (n=33), and support for vulnerable populations (n=14). Prominent themes reflecting barriers to use of RMTs included equity-related barriers (eg, affordability of technology for users, poor internet connectivity, poor health literacy; n=16), the need for quality “best practice” guidelines for use of RMTs in clinical care (n=12), and the need for additional resources to develop and support new technologies (n=11). Overall, 23 of 48 records commented on equity characteristics that stratify health opportunities and outcomes, including general characteristics that vary over time (eg, age, comorbidities; n=17), place of residence (n=11), and socioeconomic status (n=7).



**Conclusions:** Results of this rapid review highlight the breadth of RMTs being used to monitor and inform treatment of COVID-19, the potential benefits of using these technologies, and existing barriers to their use. Results can be used to prioritize further efforts in the implementation of RMTs (eg, developing “best practice” guidelines for use of RMTs and generating strategies to improve equitable access for marginalized populations).

(*JMIR Mhealth Uhealth* 2021;9(4):e24743) doi:[10.2196/24743](https://doi.org/10.2196/24743)

## KEYWORDS

remote monitoring; technology; COVID-19, telehealth; asynchronous technology; synchronous technology; mHealth; monitoring; review; barrier; benefit; equity

## Introduction

Delivering health care in the context of the COVID-19 pandemic is uniquely challenging [1]. The virus is highly contagious, potentially fatal, and symptoms that present early in the course of the illness can be mild and nonspecific [2]. Virulence of SARS-CoV-2 in health settings places health care providers and patients without COVID-19 at increased risk of contracting the illness [2]. The need to provide health care services while limiting in-person interaction places an enormous burden on already overwhelmed health systems.

Some models and technologies for delivering health care remotely existed prior to the COVID-19 pandemic. For instance, video consultation is commonly used to deliver health care services to rural and remote communities [3]. These technologies were used for a diversity of needs, including those who might otherwise have difficulty accessing expertise and health care (eg, due to remote location) [4,5]. Further, more recent public health crises of infectious diseases, including the 2002 severe acute respiratory syndrome epidemic, the 2009 H1N1 pandemic, and the 2015 Ebola epidemic, highlighted the pivotal role remote technologies can play in delivering health care [6-8]. Within a short span of a few weeks, these technologies were at the forefront and were a critical method of offering care services.

To adequately address the challenges of the COVID-19 pandemic, unprecedented deployment of remote monitoring technologies (RMTs) is underway [9,10]. Due to a lack of consensus on the definition of RMTs, we took a broad scope and included any technology that facilitates communication between the health care team and patients [11]. RMTs can be categorized as “synchronous,” where communication occurs in “real time,” such as in videoconferencing, and “asynchronous,” where there is a potential expected time delay in communication, such as email or SMS text message [11,12]. Both these forms of RMT could be delivered through a broad range of devices (eg, app on phone, wearable technology). RMTs at times reflect technology that is commonplace (eg, phone call, app on smartphone) or specialist in purpose (eg, oxygen saturation monitor, specific cell phone app). Further, RMTs have been used for a variety of purposes in health care. This has included monitoring symptoms, illness severity, and adherence to inform and deliver treatment [7].

In the context of the present pandemic, RMTs can be used when delivering health care to those with (eg, monitoring severity of illness in those with the diagnosis) or without diagnoses of COVID-19 (eg, delivering primary health care via telephone to

reduce likelihood of COVID-19 transmission). Understanding the use of RMTs for the treatment of COVID-19 is particularly pertinent in the context of an emergent pandemic and will be the focus of the present research.

In an early review of health technologies used to address COVID-19, Ming et al [9] described 110 COVID-19 asynchronous technologies including 17 that monitored patients remotely. The authors reported on the benefits of mobile phone apps, including allowing patients to self-monitor their symptoms and informing triage decisions regarding the need to access tertiary care. This use of RMTs could presumably reduce burden on hospital emergency departments. This study highlighted concerns regarding patient confidentiality and data security when using such mobile apps. Another narrative review described 17 different wearable technologies designed to remotely monitor COVID-19 signs and symptoms [10]. These devices included take-home electrocardiograms, blood pressure monitors, pulse oximeters, thermometers, and stethoscopes. These technologies allowed relevant diagnostic information to be collected. Further, use of these technologies helped optimize patient comfort and convenience while reducing the risk of COVID-19 transmission in the hospital setting and the need for hospital resources to be directed to those with milder symptoms. This review identified challenges in implementation, with a focus on technology-specific limitations—for instance, thermometer inaccuracies due to changes in ambient temperature [10].

As the pandemic develops, there will be waves of new technologies that are developed continuously to meet the emerging challenges. Some of these technologies are here to stay and could be available for future infectious diseases or be part of the routine care delivery model. Specific benefits or disadvantages of RMTs may fluctuate in relevance over time. This paper provides a snapshot that can be used as a baseline to track progress in this field over the course of the pandemic. This information will be relevant to innovators developing new RMTs for management of different health challenges, including treatment of vulnerable and marginalized patients with COVID-19.

Meeting needs of vulnerable populations requires consideration of equity in the present and for future health delivery endeavors. Equity considerations are those that focus on ways to decrease or eliminate differences in health outcomes and opportunities across groups [13]. It is vital to incorporate an equity lens since COVID-19 is more likely to impact marginalized populations [14-16]. Intentional efforts to reduce inequity are important in health care design. Considerations in inequity have been

conceptualized using the PROGRESS-Plus acronym, which is used to identify characteristics that stratify health opportunities and outcomes [17].

Thus far, published research has presented a selection of health and resource benefits and highlighted some challenges of using RMTs. However, no paper has comprehensively summarized the wide range of advantages and challenges of technologies used to deliver health care in the context of the COVID-19 pandemic. Further, there are no reviews that specifically address equity of access to health care in the context of COVID-19.

A comprehensive review and summary of RMTs, including their advantages and challenges, together with future directions for innovation or improvement, will be a valuable asset for stakeholders. This information can be used to understand the breadth of available RMT options, anticipate and prevent difficulties, and improve equitable access to health care. This is particularly relevant early in the course of a pandemic when time-sensitive information is required.

The aim of this rapid review was to identify which RMTs have been deployed to support COVID-19 health care provision early in the pandemic, and identify the barriers to and benefits of their implementation. The review will emphasize equity considerations.

## Methods

A rapid review was performed following the approach described by Cochrane Methods Group [18]. Rapid reviews offer a strategy to synthesize current knowledge in a timely manner “to meet time-sensitive decision-making needs” [1], which is valuable in contexts such as the COVID-19 global health crisis.

### Search Methods

Ovid MEDLINE and Embase were searched for relevant publications published up until July 6, 2020. The search was initially informed by a preliminary unstructured search of QxMD [19] and MEDLINE (via PubMed). The search strategy (Multimedia Appendix 1) was developed using the COVID-19 keywords from the Canadian Agency for Drugs and Technologies in Health search strings [12].

### Inclusion Criteria

Inclusion criteria for selection of publications were as follows: publications that discussed use of RMTs to deliver health care to patients with COVID-19; publications that referred to specific RMTs as opposed to general recommendations (this criterion was used to ensure barriers could be linked to a specific type of technology); and publications that were available in English, French, or Spanish.

### Exclusion Criteria

Publications were excluded if the record only described RMTs used in contact tracing or epidemiological surveillance. This criterion was used because the focus of this review was on

technologies that provide service or treatment to patients with COVID-19, rather than surveillance of the virus.

### Search Strategy

An inclusive search strategy for different types of publications was used and included editorials, reviews, and letters to the editor. This inclusive strategy was adopted considering the novelty of COVID-19 (and hence dearth of published research-based studies). Further, the inclusive search strategy reflected the potential benefits of presenting a breadth of information at this early stage of the pandemic.

### Data Collection

The primary author (EH) reviewed the title and abstract of publications then extracted key information. Key information extracted included the year, journal, design, aim of the publication, nature of the RMT, whether the RMT was synchronous or asynchronous, information gathered via RMT, equity considerations, benefits of RMT use, and barriers to RMT use.

### Data Analysis

Publications' results were synthesized using a “content analysis” approach [20]. A single reviewer (EH) developed a coding framework and mapped the RMTs used and variables monitored. Barriers and benefits of using RMT were coded inductively, and individual codes were grouped into overarching themes. A second reviewer (DOP) independently reviewed the developed coding framework against a random sample (7/48, 15%) of the publications. Further, the second reviewer coded this sample of publications against the finalized coding framework. All discrepancies (7% discrepancy rate) in the coding framework or coding were settled by consensus-based discussion.

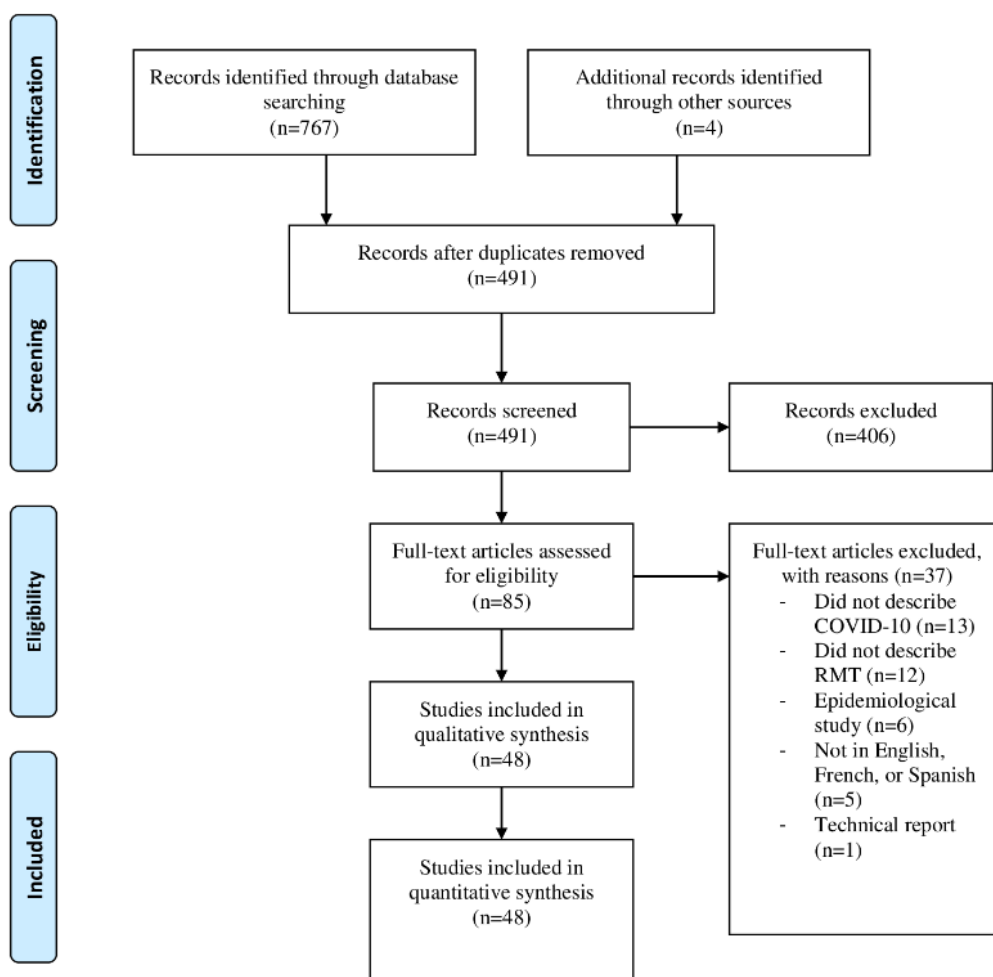
To ensure explicit consideration of health equity, relevant considerations were mapped deductively to the PROGRESS-Plus acronym [17]. The PROGRESS-Plus acronym refers to consideration of the following: place of residence, race or ethnicity or culture or language, occupation, gender or sex, religion, education, socioeconomic status, social capital, and “plus”—referring to personal characteristics associated with discrimination, features of relationships, and time-dependent relationships [17].

## Results

### Results of Database Search

Of 767 records identified through MEDLINE and Embase searches, and 4 records identified through QxMD, 491 remained after duplicates were removed (Figure 1). After title and abstract screening, 406 records were excluded from the review. Upon full-text review, 37 of 85 records were excluded because they did not include health care for patients with COVID-19 (n=13), did not describe an RMT (n=12), were epidemiological studies and/or only described contact tracing (n=6), were not in English, Spanish, or French (n=5), or were technical reports (n=1). In total, 48 publications were included in the qualitative synthesis.

**Figure 1.** PRISMA 2009 flow diagram. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RMT: remote monitoring technology.



**Characteristics of Included Publications**

The most common types of publications identified were descriptive studies or proposals describing the implementation

and development of RMTs (n=18) and reviews of multiple RMTs (n=13). Letters to the editor (n=9), editorials (n=5), one retrospective cohort study, one protocol, and one case study were also included (Table 1).

**Table 1.** Publication types and remote monitoring technologies included in the analysis.

Publication type and ID	First author [reference]	Remote monitoring technologies
<b>Implementation and development</b>		
2	Annis [21]	Mobile app
4	Bae [22]	Chat or telephone consultation, digital blood pressure monitor, digital thermometer, digital pulse oximeter
10	Faezipour [23]	Mobile app diagnostic test
13	Ford [24]	Chat or telephone or video consultation, data transmission, digital pulse oximeter, digital thermometer, electronic health record, mobile app, patient-accessed electronic health record
25	Krausz [25]	Chat or video consultation, electronic health record, mobile app, patient-accessed electronic health record, Voice over Internet Protocol technology, websites
26	Lam [12]	Video consultation
27	Lui [26]	Electronic health record, mobile apps
28	Mann [27]	Electronic health record, video consultation
30	Medina [28]	E-prescription, mobile app, telephone consultation
32	Naik [29]	Closed-circuit television cameras, digital blood pressure monitor, digital electrocardiogram or heart rate monitor, digital end-tidal carbon dioxide monitor, digital pulse oximeter, digital respiratory rate monitor, mobile device video
34	Petrocelli [30]	Home diagnostic test
35	Saleem [31]	Automated SMS text messaging program
36	Schinköthe [32]	Electronic health record, mobile app
39	Song [33]	Video consultation
40	Sossai [34]	Chat or video consultation, mobile app
42	Timmers [35]	Mobile app
44	Vaira [36]	Home diagnostic test, telephone consultation
<b>Implementation and development proposal</b>		
38	Sharma [37]	Wearable devices or biometric clothing
<b>Retrospective cohort study</b>		
47	Xu [38]	Chat or telephone consultation
<b>Case study</b>		
20	Huang [39]	Chat consultation, mobile app
<b>Review</b>		
1	Alwashimi [40]	Artificial intelligence-enabled request for assistance, chatbots, digital stethoscope, drone-delivered home diagnostic tests, glucometer, mobile app smart inhaler, video consultation, wearable devices or biometric clothing, website
7	Crawford [41]	Video consultation
8	Ding [10]	Digital electrocardiogram or heart rate monitor, digital blood pressure monitor, digital respiratory rate monitor, digital pulse oximeter, digital stethoscope, digital thermometer, mobile device recording, wearable devices or biometric clothing
11	Fagherazzi [42]	Chat or telephone or video consultation, e-prescription
18	Hong [43]	E-prescription, mobile phone app, telephone consultation, video consultation
19	Horowitz [44]	Mobile app diagnostic test, mobile app
23	Kannampallil [45]	Mobile app, wearable devices or biometric clothing
24	Keshvardoost [46]	Electronic health record, telephone or video consultation
29	Massaroni [47]	Laptop camera, mobile device camera, radar/Wi-Fi transmitter-receiver, smart mattress, tablet camera, wearable devices or biometric clothing
31	Ming [48]	Mobile app

Publication type and ID	First author [reference]	Remote monitoring technologies
37	Sharma [49]	Digital blood pressure monitor, digital electrocardiogram or heart rate monitor, digital glucometer, digital pulse oximeter, digital posture, digital respiratory rate monitor, digital temperature monitor, digital weight monitor, patient-accessed electronic health record, video consultation
41	Thulesius [50]	Video consultation
45	Watson [51]	Home diagnostic test, mobile app, video consultation
48	Ye [52]	Telephone or video consultation
<b>Editorial</b>		
6	Cohen [53]	Tablet app, telephone consultation
9	Edelman [54]	Electronic health record, video consultation
16	Grenngalgh [55]	Video consultation
22	John [56]	Video consultation
<b>Protocol</b>		
15	Greenhalgh [57]	Telephone or video consultation
<b>Letter to the editor</b>		
3	Anonymous [58]	Telephone consultation
5	Barsom [59]	Electronic health records, patient-accessed electronic health record, tablet, video consultation
12	Fitz [60]	Chat consultation
14	Giansanti [61]	Mobile app, video consultation
17	Hau [62]	Telephone or video consultation
21	Jamil [63]	Telephone or video consultation
33	Nair [64]	Electronic health record mobile app, telephone or video consultation
43	Trethewey [65]	Telephone or video consultation
46	Wei [66]	Video consultation

## Description of RMTs

These publications identified 35 distinct types of RMTs (32 asynchronous and 3 synchronous; [Table 2](#)). Overall, 35 of the 48 studies identified synchronous technologies, including video (n=25), telephone (n=15), chat (n=8), and mobile or cellular device video (n=3) consultations. In total, 33 studies reported asynchronous technologies, including mobile or cellular apps for patients to manually enter their symptoms or education (n=17), home diagnostic tests (n=6), remote access to electronic health records by provider (n=8), digital pulse oximeters (n=5),

biometric clothing (n=5), and remote access to electronic health records by patient (n=5). Data were collected through manual input by patients, caregivers, and health care practitioners (n=45; [Multimedia Appendix 2](#)), and automatically through the use of biosensor technology (n=22; [Multimedia Appendix 3](#)). The most commonly reported manually inputted outcome measures were “symptoms” (n=18), dyspnea or shortness of breath (n=7), pulse oximetry (n=7), mental health (n=7), temperature (n=6), and presence or absence of fever (n=6). The most frequently reported biosensor measures were temperature (n=5) and cardiac activity (as measured by an electrocardiogram; n=5).

**Table 2.** Summary of remote monitoring technologies grouped into synchronous or asynchronous.

Type of remote monitoring technology	Total publications, n
<b>Asynchronous</b>	
Mobile or cellular app	17
Electronic health record (eg, cloud based)	8
Home diagnostic tests	6
Digital (eg, Bluetooth) pulse oximeter	5
Wearable device or biometric clothing	5
Patient-accessed electronic health record	5
Digital thermometer	4
Digital blood pressure sensor	4
Mobile or cellular app diagnostic test (acoustic signal, olfactory and gustatory test)	3
Digital heart rate monitor, respiratory rate monitor, and electrocardiogram (eg, measure heart rate variability)	3
Tablet (eg, iPads) may be used for video	3
E-prescription	3
Website to monitor symptoms	2
Digital stethoscope	2
Glucometer	2
Data transmission technology	2
Automated SMS text messaging program	1
Chatbot, drone, artificial intelligence-enabled request for assistant, smart inhaler	1
Closed-circuit television camera and digital end-tidal carbon dioxide monitor	1
Digital weight and posture sensors	1
Smart mattress	1
Voice over Internet Protocol technology	1
<b>Synchronous</b>	
Video consultation	25
Telephone consultation	15
Chat consultation	7
Mobile or cellular device video consultation	2

### Benefits of RMTs

Almost all publications (46/48, 96%) provided some description of the reported benefits of RMTs. In total, 34 distinct codes were developed inductively from these documents ([Multimedia Appendix 4](#)). These were grouped into 10 themes ([Table 3](#)): “reduces infection risk” (n=33 articles), “reduces burden of care” (eg, limits hospital beds used; n=28), “supports vulnerable populations” (n=14), “reduces costs” (n=12), “improves patient experience” (n=11), “promotes knowledge development” (n=8), “facilitates navigation through health care system” (n=8), “improves health outcomes” (n=6), “supports public health initiatives” (n=5), and “technology-specific benefits” (n=4).

The most common theme, “reduces infection risk,” included 3 codes: “reduces risk of transmission generally” (n=18), “reduces exposure of health care practitioners” (n=11), and “reduces cross-contamination or clustering” (n=10). A total of 28 codes

were mapped to the theme “reduces burden of care,” notably that RMTs “reduce the burden on hospitals” (eg, limit hospital beds used; n=17), “provide continuous accessible data or monitoring” (n=12), and “reduce burden of time on health care workers” (eg, using asynchronous technology or reducing burden overall; n=6). In addition, 5 codes were mapped to the theme “RMTs support vulnerable populations,” including “reduces need for transfer of vulnerable patients” (n=6) and “supports patient mental health” (eg, relieves stress, supports anxious patients; n=5). A total of 3 codes were mapped to the theme “reduces cost,” notably “reduces health care system or public health agency costs” (n=12). A total of 3 codes were mapped to the theme “improves patient experience,” including “improves patient initiative, engagement, autonomy, or self-management” (n=5). In addition, two codes were mapped to “facilitates navigation through the health care system,” notably “facilitated follow-up, continuity of care, or linkage to care” (n=5). The theme “improves health outcomes” contained

only one code, “provides rapid identification of infection or clinical deterioration for timely treatment of COVID-19” (n=6). Finally, the theme “supports public health initiatives” contained

2 codes, including “delivers educational messages” (eg, fights disinformation or the “infodemic”; n=5). All other codes were noted by <5 publications.

**Table 3.** Perceived benefits of remote monitoring technologies for patients with COVID-19.

Themes	Total publications, n
Reduces infection risk	33
Reduces burden of care	28
Supports vulnerable populations	14
Reduces costs	12
Improves patient experience	11
Promotes knowledge development	8
Facilitates navigation through health care system	8
Improves health outcomes	6
Supports public health initiatives	5
Technology-specific benefits	4

### Barriers to Using RMTs

Many publications (38/48, 80%) reported barriers, challenges, and/or concerns regarding implementation of RMTs (Table 4). In all, 54 distinct codes were developed inductively (Multimedia Appendix 5) and mapped to 15 themes: “equity-related barriers” (n=16), “a lack of RMT implementation guidelines and research” (n=12), “resources required for technology development and implementation” (n=11), “challenging patient experiences of RMTs” (n=10), “confidentiality-related concerns” (n=10), “workforce training” (n=8), “quality of information” (n=8), “communication-related barriers” (n=7), “ethical concerns with RMTs” (n=7), “policy requirements” (n=7), “quality of care concerns” (n=4), “technology-specific barriers” (n=3), “technology integration-related barriers” (n=2), and “financial barriers” (n=2). A total of 8 codes were mapped to the theme “equity-related barriers,” including “lack of access to RMTs in low-resource settings” (eg, patients experiencing homelessness, neighborhoods without access to libraries, households without internet or devices, low-income communities unable to afford RMTs or share RMTs; n=8), “low network quality or internet connectivity or bandwidth,” which can impact quality of care (n=6), and “low patient health literacy” (n=5). A total of 5 codes were mapped to the theme “a lack of RMT implementation

guidelines and research,” with 9 publications reporting a “paucity of high-quality data or guidelines to support effective and safe RMT, particularly in acute care.” A total of 7 codes were mapped to the theme “resources required for technology development and implementation,” including “inadequate control of patient flow with some RMTs” (eg, the fluctuating recruitment of patients should be matched with staffing; n=5). A total of 4 codes were mapped to the theme “challenging patient experiences of RMTs,” including “the complexity or intrusiveness of switching to online consultation or remote monitoring and disruption to patient or worker processes and routines” (n=5). A total of 3 codes were mapped to the theme “confidentiality-related barriers,” including “the need to address privacy concern when implementing RMTs” (n=9). A total of 3 codes were mapped to the theme “workforce training,” including “the need for additional workforce education and training in use of RMTs” (n=6). A total of 2 codes were mapped to the theme “quality of information,” including “issues regarding the quality of health information reported or collected” (eg, self-reporting; n=6). Finally, 3 codes were mapped to the theme “ethical concerns with RMTs,” of which the majority (n=5) of the publications reported general ethical concerns. All other codes were reported by <5 publications.

**Table 4.** Perceived barriers to using remote monitoring technologies for patients with COVID-19.

Themes	Total publications, n
Equity-related barriers	16
A lack of remote monitoring technology implementation guidelines and research	12
Resources required for technology development and implementation	11
Challenging patient experiences of remote monitoring technologies	11
Confidentiality-related barriers	10
Workforce training	8
Quality of information	8
Communication-related barriers	7
Ethical concerns with remote monitoring technologies	7
Policy requirements	7
Quality of care	4
Technology-specific barriers	3
Technology integration-related barriers	2
Financial barriers	2

### Equity Factors

Equity groups were mapped deductively to the PROGRESS-Plus acronym (Table 5). Several publications (n=23) noted how remote monitoring provided information on characteristics of populations that could stratify health opportunities and outcomes (Multimedia Appendix 6). The most frequently reported characteristic was the “Plus” code (n=17). This “Plus” code refers to specific patient populations (ie, patients who distrust the health care system, those with chronic conditions or comorbidities, immune-suppressed patients, pregnant women, acutely unwell patients, patients with cognitive impairment, nonadherent patients, seniors, and youth). The next most common PROGRESS-Plus code was “place of residence.” A total of 11 publications commented on patient place of residence

as impacting the implementation of RMTs, including rural and remote residences, nursing homes, or homelessness. A total of 7 publications reported socioeconomic factors, such as a lack of community funding in low- and middle-income countries, lack of access to technology, or low-income patients. A total of 3 publications reported on race or ethnicity or culture or language, occupation, and gender or sex, respectively. Occupations considered were frontline workers and veterans. Finally, in terms of PROGRESS-Plus equity considerations, 2 publications commented on education, specifically the health literacy level of patients, which could impact their ability to understand infection control information or assess the quality of unregulated health information. Only one publication commented on social capital and none on religion.



**Table 5.** PROGRESS-Plus themes.

Theme	Total publications, n	Demonstrative quote
Place of residence	11	“It is also important to consider that some countries may not have the technological infrastructure to support [digital health]. Furthermore, there will be a significant proportion of the population who will not have access to technology or internet connectivity” [40]
Race or ethnicity or culture or language	3	“Mounting evidence suggests that the COVID-19 pandemic has far greater associated morbidity and mortality in racialized groups that struggle with poverty and poor access to health care; the pandemic has also been suggested to compound pre-existing inequities. Similarly, there has been a lack of attention to health equity in the development of digital health solutions” [41]
Occupation	3	“Individuals exposed to the public (such as transit workers and police/force/emergency medical services (EMS) workers) and cultural settings with risk for infection (eg, multi-family settings with multiple house members working with high risk for COVID-19 exposure settings), in addition to those in meat packing plants or the front-line grocery store workers, are especially in the high-risk exposure category. This additionally emphasizes the need for developing testing models of the breathing app from the breathing sound database.” [23]
Gender or sex	3	“In Dover and Belon’s model, which informs the foundation of the [Digital Health Equity Framework], the process of social stratification within economic and cultural social contexts refers to the hierarchical allocation and unequal distribution of power, prestige, and resources; this stratification assigns individuals to a social location, which is defined by intersectional factors such as race, age, income, geography, rurality, gender, ability, and occupation as well as other social factors” [41]
Religion	0	— <sup>a</sup>
Education	2	“Our findings suggest that even in high-income countries, such as the [United Arab Emirates] with modern digital access and high general literacy rates, health literacy may pose an obstacle for the adoption of telemedicine. It is, therefore, critical for countries worldwide to improve health literacy to optimize patient access and engagement in this expanding world of digital medical care delivery.” [64]
Socioeconomic status	7	“Lastly, some developing countries face major obstacles to the effective delivery of digital health solutions in rural and remote locations, such as incomplete or insufficient basic digital infrastructures (eg, computers, internet networks, and electricity), lack of sustainable funding to develop, operate, and maintain digital platforms, and high telecommunication costs” [42]
Social capital	1	See “Gender or sex” above [41]
Plus	17	“The network has focused on groups particularly vulnerable to severe symptoms of COVID-19, including the elderly, pregnant women, children, and patients with chronic health problems.” [43]

<sup>a</sup>Not available.

## Discussion

### Principal Findings

A rapid review was conducted to present the different types of RMTs deployed in response to the COVID-19 pandemic, the perceived benefits of these technologies, and the perceived barriers or challenges to their use in clinical work. There was special emphasis placed on equity considerations.

Results indicate that many different RMTs are used when providing health care to patients with COVID-19. Most RMTs mirrored the real-time interpersonal communication processes involved in pre-pandemic care, except that the care was being delivered with physical distance—for instance, the use of videoconferencing (25/48, 52%) and teleconferencing (15/48, 31%) rather than in-person appointments, and accessing medical records remotely (8/48, 17%) rather in the office. Video was likely favored because of the substantial advantage of being able to use nonverbal cues and assessment. These innovations

were pertinent to monitoring, triage, and diagnosis (eg, symptom tracking using mobile apps).

### Benefits of RMTs and Implications

The most common reported benefit of RMTs was that it reduced the risk of transmission of COVID-19 (n=18), including to health care practitioners (n=10). This is unsurprising, considering that concern about managing infection rates is one of the forces driving rapid implementation of RMTs across many specialties [67]. Further, with increasing pressure on the health care system, RMTs could aid in reducing the burden on hospitals (n=17), health care system costs (n=10), and the burden of time of health care practitioners (n=16).

Use of these RMTs began or substantially increased in response to COVID-19. If perceived advantages are borne out, this may lead to wider adoption beyond the pandemic.

There may be several implications when considering the breadth of RMTs already in use. First, both health care providers and

patients may become accustomed to the remote modality and its benefits, including increased convenience, meaning that RMT use could continue beyond the duration of the pandemic. Health care during the pandemic has increased patient involvement in their own care (eg, symptom monitoring). This may lead to a permanent shift in health care culture, in which patients' active participation is incorporated into the health provision of other conditions. The development of health-related software for signs and symptom monitoring and triage could also continue and become refined and expand the tools through which quality health care is delivered. When considering whether or which changes might endure, it might be helpful to distinguish between advantages of RMTs that are only relevant to COVID-19 infection risk and those where the benefits might persist beyond COVID-19 (eg, widespread use of health-related apps to better monitor symptoms for other conditions). It is possible and likely that the processes surrounding their use will be refined to address and reduce concerns and barriers.

### Barriers to Use of RMTs

The most commonly reported barrier themes were a lack of guidance (n=12) and increased resources needed (n=11) for implementation, development, and use of RMTs to treat COVID-19. These barriers are likely due to the novel nature of COVID-19 and thus could become less relevant over time. Another main concern cited by several publications was that rigorous privacy and security settings would be necessary to protect patient information (n=9). However, despite emphasizing the importance of privacy and security, only 5 publications of 18 describing implementation of a specific RMT reported on security and privacy features or policies of the software used [24-26,32,35]. Additionally, two publications noted that use of RMTs could break down the humanitarian core of care as well as patient-provider communication [43,50]. One way to promote effective patient communication is to design user-friendly technology with two-way communication [68]. Although patient involvement in technology development can be used to effectively tailor the technology to the specific needs of the patient [69], no publications reported this in practice. This is unfortunate considering the impact it can have on the success of an RMT [70]. Future technologies should involve rigorous user evaluation based on feedback from patients. The extent of patient involvement in RMT implementation should be thoroughly described to support use of the technology. Lastly, it will take time and resources to bring RMTs to scale; information regarding clinical utility and cost will help ascertain which should be prioritized for investment of resources to aid in this development.

### Equity Factors

Equity factors also proved important to consider when implementing RMTs. Health interventions should be tailored according to population needs or they risk increasing health inequities [17]. The most common overall barrier theme was concerns regarding equitable use of RMTs (n=15). This is concerning considering marginalized groups are already disproportionately impacted by COVID-19. There is a higher incidence of infection, as well as poorer outcomes, among racialized communities and ethnic minorities [71-73]. More

than half of the included publications (n=28) noted challenges that might be faced by minorities or emphasized the importance of training health practitioners in equitable digital health implementation and supporting low-income communities. Alternatively, possible benefits were reported in some papers (n=7), including the ability to support vulnerable populations—for example, medically vulnerable individuals, or those living in rural or remote communities. Overall, there appeared to be conflicting perspectives on whether RMTs would decrease or increase health inequities. There is evidence that intervention-generated inequities—those caused by use of interventions that provide limited benefit to vulnerable populations—can decrease health outcomes among marginalized groups [13]. It is important to track characteristics that may stratify health outcomes in order to assess which RMTs will benefit vulnerable populations. Overall, 23 of 48 total papers (48%) and 7 of 18 (39%) implementation and development papers reported on characteristics of populations that stratify health outcomes. It is critical that an increasing number of publications start to analyze data based on characteristics that influence equitable outcomes to better support vulnerable groups and reduce health inequities.

### Limitations

This work has several time-related limitations due to the rapid nature of this review. Only one screener (EH) reviewed the titles, abstracts, and full texts of the records and coded during the selection process. Further, given the delay between the use of an intervention and publication of its evaluation, the technologies presented here only reflect what was used during a specific window of time and as new technologies are rapidly developed and implemented, there will be a need to reassess new and emerging benefits of and barriers to patient care. Further periodic reviews should be conducted to assess how the use of RMTs evolves. In addition, the quality of the analysis was limited. Furthermore, the frequency of reporting does not necessarily equal the magnitude of importance or the impact on outcomes for patients. Lastly, certain papers reported on multiple technologies but did not provide benefits and barriers specific to each RMT individually, which limits the generalizability of our findings. As such, most benefits and barriers reported in this paper were categorized generally as applicable to both synchronous and asynchronous RMTs. Undoubtedly, many of these benefits and barriers will be more or less relevant depending on the technology (eg, reduced isolation is more applicable for treatment via video consultation than for use of a digital pulse oximeter). Future reviews should focus on separating benefits and barriers according to technology type if possible.

### Conclusion

This rapid review summarizes RMTs that were used in the early stages of the COVID-19 pandemic and provides insights on the benefits these technologies could provide for future use. It also highlights perceived barriers to implementation of RMTs that should be addressed in ongoing development projects. Guidelines and policies developed for implementing RMTs ought to be mindful of the identified barriers. RMTs have an important role to play in supporting patients and communities

through these unprecedented times. One key recommendation is to establish best practices in the development of RMTs so they are both equitable and effective going forward.

---

## Acknowledgments

BL received the Canadian Institutes for Health Research (CIHR), Strategy for Patient-Oriented Research Mentorship Chair in Innovative Clinical Trials for HIV Care. KM, a postdoctoral fellow, is funded by this mentorship chair. BL is also supported by a career award LE 250 from Quebec's Ministry of Health for researchers in Family Medicine.

Our work is one of the critical research programs being supported by the McGill Interdisciplinary Initiative in Infection and Immunity (MI4), with seed funding from the McGill University Health Centre Foundation.

This work is supported by CIHR funding awarded to MPP and BL ("Real time evaluation of the deployment of connected technologies and of the partnership of services and care during the COVID-19 sanitary crisis – the Techno-COVID-Partnership program," 2020-05-12 Operating Grant: COVID-19 May 2020 Rapid Research Funding Opportunity - grant 444186 #).

---

## Authors' Contributions

EH, KM, and BL all contributed to the conceptualization of this project. EH developed the search strategy, completed title and abstract screening and full-text review. EH collected and analyzed the data from included publications, which were verified by DOP. EH prepared the written manuscript. KM, KE, DOP, MPP, JC, TH, and BL reviewed the manuscript and provided feedback.

---

## Conflicts of Interest

Author EH is employed by author BL. Author TH is the cofounder of Opal, a patient portal and app. The other authors have no conflicts to declare.

---

### Multimedia Appendix 1

COVID-19 remote monitoring search strategy: July 6, 2020.

[\[DOCX File , 17 KB - mhealth\\_v9i4e24743\\_app1.docx \]](#)

---

### Multimedia Appendix 2

Physiological parameters and symptoms that are manually inputted into remote monitoring technologies.

[\[DOCX File , 17 KB - mhealth\\_v9i4e24743\\_app2.docx \]](#)

---

### Multimedia Appendix 3

IDs of records that reported automatically recorded biosensor outcome measures.

[\[DOCX File , 15 KB - mhealth\\_v9i4e24743\\_app3.docx \]](#)

---

### Multimedia Appendix 4

IDs of records that reported benefits of remote monitoring technologies.

[\[DOCX File , 18 KB - mhealth\\_v9i4e24743\\_app4.docx \]](#)

---

### Multimedia Appendix 5

IDs of records that reported barriers to use of remote monitoring technologies.

[\[DOCX File , 25 KB - mhealth\\_v9i4e24743\\_app5.docx \]](#)

---

### Multimedia Appendix 6

IDs of records that identified characteristics that stratify health opportunities or outcomes mapped to the PROGRESS-Plus framework.

[\[DOCX File , 13 KB - mhealth\\_v9i4e24743\\_app6.docx \]](#)

---

## References

1. Hartling L, Guise J, Kato E, Anderson J, Belinson S, Berliner E, et al. A taxonomy of rapid reviews links report types and methods to specific decision-making contexts. *J Clin Epidemiol* 2015 Dec;68(12):1451-62.e3. [doi: [10.1016/j.jclinepi.2015.05.036](https://doi.org/10.1016/j.jclinepi.2015.05.036)] [Medline: [26278023](https://pubmed.ncbi.nlm.nih.gov/26278023/)]
2. Esakandari H, Nabi-Afjadi M, Fakkari-Afjadi J, Farahmandian N, Miresmaeili S, Bahreini E. A comprehensive review of COVID-19 characteristics. *Biol Proced Online* 2020 Aug 4;22(1):19 [FREE Full text] [doi: [10.1186/s12575-020-00128-2](https://doi.org/10.1186/s12575-020-00128-2)] [Medline: [32774178](https://pubmed.ncbi.nlm.nih.gov/32774178/)]

3. Pignatiello A, Teshima J, Boydell KM, Minden D, Volpe T, Braunberger PG. Child and youth telepsychiatry in rural and remote primary care. *Child Adolesc Psychiatr Clin N Am* 2011 Jan;20(1):13-28. [doi: [10.1016/j.chc.2010.08.008](https://doi.org/10.1016/j.chc.2010.08.008)] [Medline: [21092909](https://pubmed.ncbi.nlm.nih.gov/21092909/)]
4. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. *Telemed J E Health* 2017 Jan;23(1):3-17 [FREE Full text] [doi: [10.1089/tmj.2016.0051](https://doi.org/10.1089/tmj.2016.0051)] [Medline: [27116181](https://pubmed.ncbi.nlm.nih.gov/27116181/)]
5. Queirós A, Alvarelhão J, Cerqueira M, Silva A, Santos M, Pacheco Rocha N. Remote Care Technology: A Systematic Review of Reviews and Meta-Analyses. *Technologies* 2018 Feb 10;6(1):22. [doi: [10.3390/technologies6010022](https://doi.org/10.3390/technologies6010022)]
6. Anderson RM, Fraser C, Ghani AC, Donnelly CA, Riley S, Ferguson NM, et al. Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. *Philos Trans R Soc Lond B Biol Sci* 2004 Jul 29;359(1447):1091-1105 [FREE Full text] [doi: [10.1098/rstb.2004.1490](https://doi.org/10.1098/rstb.2004.1490)] [Medline: [15306395](https://pubmed.ncbi.nlm.nih.gov/15306395/)]
7. Ohannessian R. Telemedicine: Potential applications in epidemic situations. *European Research in Telemedicine / La Recherche Européenne en Télémedecine* 2015 Sep;4(3):95-98. [doi: [10.1016/j.eurtele.2015.08.002](https://doi.org/10.1016/j.eurtele.2015.08.002)]
8. Tom-Aba D, Nguku PM, Arinze CC, Krause G. Assessing the Concepts and Designs of 58 Mobile Apps for the Management of the 2014-2015 West Africa Ebola Outbreak: Systematic Review. *JMIR Public Health Surveill* 2018 Oct 29;4(4):e68 [FREE Full text] [doi: [10.2196/publichealth.9015](https://doi.org/10.2196/publichealth.9015)] [Medline: [30373727](https://pubmed.ncbi.nlm.nih.gov/30373727/)]
9. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile Health Apps on COVID-19 Launched in the Early Days of the Pandemic: Content Analysis and Review. *JMIR mHealth uHealth* 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: [10.2196/19796](https://doi.org/10.2196/19796)] [Medline: [32609622](https://pubmed.ncbi.nlm.nih.gov/32609622/)]
10. Ding X, Clifton D, Ji N, Lovell NH, Bonato P, Chen W, et al. Wearable Sensing and Telehealth Technology with Potential Applications in the Coronavirus Pandemic. *IEEE Rev Biomed Eng* 2021;14:48-70. [doi: [10.1109/rbme.2020.2992838](https://doi.org/10.1109/rbme.2020.2992838)]
11. Verhoeven F, Tanja-Dijkstra K, Nijland N, Eysenbach G, van Gemert-Pijnen L. Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. *J Diabetes Sci Technol* 2010 May 01;4(3):666-684 [FREE Full text] [doi: [10.1177/193229681000400323](https://doi.org/10.1177/193229681000400323)] [Medline: [20513335](https://pubmed.ncbi.nlm.nih.gov/20513335/)]
12. Lam PW, Sehgal P, Andany N, Mubareka S, Simor AE, Ozaldin O, et al. A virtual care program for outpatients diagnosed with COVID-19: a feasibility study. *CMAJ Open* 2020 May 23;8(2):E407-E413 [FREE Full text] [doi: [10.9778/cmajo.20200069](https://doi.org/10.9778/cmajo.20200069)] [Medline: [32447283](https://pubmed.ncbi.nlm.nih.gov/32447283/)]
13. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health* 2013 Feb 08;67(2):190-193. [doi: [10.1136/jech-2012-201257](https://doi.org/10.1136/jech-2012-201257)] [Medline: [22875078](https://pubmed.ncbi.nlm.nih.gov/22875078/)]
14. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: Systematic review and thematic synthesis of qualitative studies. *Int J Med Inform* 2019 Apr;124:78-85 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.01.013](https://doi.org/10.1016/j.ijmedinf.2019.01.013)] [Medline: [30784430](https://pubmed.ncbi.nlm.nih.gov/30784430/)]
15. Crawford A, Serhal E. Digital Health Equity and COVID-19: The Innovation Curve Cannot Reinforce the Social Gradient of Health. *J Med Internet Res* 2020 Jun 02;22(6):e19361 [FREE Full text] [doi: [10.2196/19361](https://doi.org/10.2196/19361)] [Medline: [32452816](https://pubmed.ncbi.nlm.nih.gov/32452816/)]
16. Pandian PS. An Overview of Telemedicine Technologies for Healthcare Applications. In: *Wearable Technologies: Concepts, Methodologies, Tools, and Applications*. Hershey, PA, USA: IGI Global; 2018:732-759.
17. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014 Jan;67(1):56-64. [doi: [10.1016/j.jclinepi.2013.08.005](https://doi.org/10.1016/j.jclinepi.2013.08.005)] [Medline: [24189091](https://pubmed.ncbi.nlm.nih.gov/24189091/)]
18. Garrity C, Gartlehner G, Nussbaumer-Streit B, King VJ, Hamel C, Kamel C, et al. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. *J Clin Epidemiol* 2021 Feb;130:13-22 [FREE Full text] [doi: [10.1016/j.jclinepi.2020.10.007](https://doi.org/10.1016/j.jclinepi.2020.10.007)] [Medline: [33068715](https://pubmed.ncbi.nlm.nih.gov/33068715/)]
19. QxMD. URL: <https://qxmd.com/> [accessed 2021-02-21]
20. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008 Apr;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
21. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson JA, Billecke S, et al. Rapid implementation of a COVID-19 remote patient monitoring program. *J Am Med Inform Assoc* 2020 Aug 01;27(8):1326-1330 [FREE Full text] [doi: [10.1093/jamia/ocaa097](https://doi.org/10.1093/jamia/ocaa097)] [Medline: [32392280](https://pubmed.ncbi.nlm.nih.gov/32392280/)]
22. Bae YS, Kim KH, Choi SW, Ko T, Jeong CW, Cho B, et al. Information Technology-Based Management of Clinically Healthy COVID-19 Patients: Lessons From a Living and Treatment Support Center Operated by Seoul National University Hospital. *J Med Internet Res* 2020 Jun 12;22(6):e19938 [FREE Full text] [doi: [10.2196/19938](https://doi.org/10.2196/19938)] [Medline: [32490843](https://pubmed.ncbi.nlm.nih.gov/32490843/)]
23. Faezipour M, Abuzneid A. Smartphone-Based Self-Testing of COVID-19 Using Breathing Sounds. *Telemed J E Health* 2020 Oct 02;26(10):1202-1205. [doi: [10.1089/tmj.2020.0114](https://doi.org/10.1089/tmj.2020.0114)] [Medline: [32487005](https://pubmed.ncbi.nlm.nih.gov/32487005/)]
24. Ford D, Harvey J, McElligott J, King K, Simpson KN, Valenta S, et al. Leveraging health system telehealth and informatics infrastructure to create a continuum of services for COVID-19 screening, testing, and treatment. *J Am Med Inform Assoc* 2020 Dec 09;27(12):1871-1877 [FREE Full text] [doi: [10.1093/jamia/ocaa157](https://doi.org/10.1093/jamia/ocaa157)] [Medline: [32602884](https://pubmed.ncbi.nlm.nih.gov/32602884/)]
25. Krausz M, Westenberg JN, Vigo D, Spence RT, Ramsey D. Emergency Response to COVID-19 in Canada: Platform Development and Implementation for eHealth in Crisis Management. *JMIR Public Health Surveill* 2020 May 15;6(2):e18995 [FREE Full text] [doi: [10.2196/18995](https://doi.org/10.2196/18995)] [Medline: [32401218](https://pubmed.ncbi.nlm.nih.gov/32401218/)]

26. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, et al. A COVID-19 Risk Assessment Decision Support System for General Practitioners: Design and Development Study. *J Med Internet Res* 2020 Jun 29;22(6):e19786 [FREE Full text] [doi: [10.2196/19786](https://doi.org/10.2196/19786)] [Medline: [32540845](https://pubmed.ncbi.nlm.nih.gov/32540845/)]
27. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: Evidence from the field. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
28. Medina M, Babiuch C, Card M, Gavrilescu R, Zafirau W, Boose E, et al. Home monitoring for COVID-19. *Cleve Clin J Med* 2020 Jun 11;87(5):1-4 [FREE Full text] [doi: [10.3949/ccjm.87a.ccc028](https://doi.org/10.3949/ccjm.87a.ccc028)] [Medline: [32409432](https://pubmed.ncbi.nlm.nih.gov/32409432/)]
29. Naik BN, Gupta R, Singh A, Soni SL, Puri GD. Real-Time Smart Patient Monitoring and Assessment Amid COVID-19 Pandemic - an Alternative Approach to Remote Monitoring. *J Med Syst* 2020 Jun 13;44(7):131 [FREE Full text] [doi: [10.1007/s10916-020-01599-2](https://doi.org/10.1007/s10916-020-01599-2)] [Medline: [32533379](https://pubmed.ncbi.nlm.nih.gov/32533379/)]
30. Petrocelli M, Ruggiero F, Baietti AM, Pandolfi P, Salzano G, Salzano FA, et al. Remote psychophysical evaluation of olfactory and gustatory functions in early-stage coronavirus disease 2019 patients: the Bologna experience of 300 cases. *J Laryngol Otol* 2020 Jul 28;134(7):571-576. [doi: [10.1017/s0022215120001358](https://doi.org/10.1017/s0022215120001358)]
31. Saleem J, Read J, Loehr B, Frisbee KL, Wilck NR, Murphy JJ, et al. Veterans' response to an automated text messaging protocol during the COVID-19 pandemic. *J Am Med Inform Assoc* 2020 Aug 01;27(8):1300-1305 [FREE Full text] [doi: [10.1093/jamia/ocaa122](https://doi.org/10.1093/jamia/ocaa122)] [Medline: [32470974](https://pubmed.ncbi.nlm.nih.gov/32470974/)]
32. Schinköthe T, Gabri MR, Mitterer M, Gouveia P, Heinemann V, Harbeck N, et al. A Web- and App-Based Connected Care Solution for COVID-19 In- and Outpatient Care: Qualitative Study and Application Development. *JMIR Public Health Surveill* 2020 Jun 01;6(2):e19033 [FREE Full text] [doi: [10.2196/19033](https://doi.org/10.2196/19033)] [Medline: [32406855](https://pubmed.ncbi.nlm.nih.gov/32406855/)]
33. Song X, Liu X, Wang C. The role of telemedicine during the COVID-19 epidemic in China-experience from Shandong province. *Crit Care* 2020 Apr 28;24(1):178 [FREE Full text] [doi: [10.1186/s13054-020-02884-9](https://doi.org/10.1186/s13054-020-02884-9)] [Medline: [32345359](https://pubmed.ncbi.nlm.nih.gov/32345359/)]
34. Sossai P, Uguccioni S, Casagrande S. Telemedicine and the 2019 coronavirus (SARS-CoV-2). *Int J Clin Pract* 2020 Oct 08;74(10):e13592 [FREE Full text] [doi: [10.1111/ijcp.13592](https://doi.org/10.1111/ijcp.13592)] [Medline: [32563197](https://pubmed.ncbi.nlm.nih.gov/32563197/)]
35. Timmers T, Janssen L, Stohr J, Murk J, Berrevoets M. Using eHealth to Support COVID-19 Education, Self-Assessment, and Symptom Monitoring in the Netherlands: Observational Study. *JMIR Mhealth Uhealth* 2020 Jun 23;8(6):e19822 [FREE Full text] [doi: [10.2196/19822](https://doi.org/10.2196/19822)] [Medline: [32516750](https://pubmed.ncbi.nlm.nih.gov/32516750/)]
36. Vaira LA, Salzano G, Petrocelli M, Deiana G, Salzano FA, De Riu G. Validation of a self-administered olfactory and gustatory test for the remotely evaluation of COVID-19 patients in home quarantine. *Head Neck* 2020 Jul 09;42(7):1570-1576 [FREE Full text] [doi: [10.1002/hed.26228](https://doi.org/10.1002/hed.26228)] [Medline: [32357379](https://pubmed.ncbi.nlm.nih.gov/32357379/)]
37. Sharma M, Sharma S, Singh G. Remote monitoring of physical and mental state of 2019-nCoV victims using social internet of things, fog and soft computing techniques. *Comput Methods Programs Biomed* 2020 Nov;196:105609. [doi: [10.1016/j.cmpb.2020.105609](https://doi.org/10.1016/j.cmpb.2020.105609)] [Medline: [32593062](https://pubmed.ncbi.nlm.nih.gov/32593062/)]
38. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and Management of Home-Quarantined Patients With COVID-19 Using a WeChat-Based Telemedicine System: Retrospective Cohort Study. *J Med Internet Res* 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: [10.2196/19514](https://doi.org/10.2196/19514)] [Medline: [32568727](https://pubmed.ncbi.nlm.nih.gov/32568727/)]
39. Huang S, Xiao Y, Yan L, Deng J, He M, Lu J, et al. Implications for Online Management: Two Cases with COVID-19. *Telemed J E Health* 2020 Apr;26(4):487-494. [doi: [10.1089/tmj.2020.0066](https://doi.org/10.1089/tmj.2020.0066)] [Medline: [32233973](https://pubmed.ncbi.nlm.nih.gov/32233973/)]
40. Alwashmi MF. The Use of Digital Health in the Detection and Management of COVID-19. *Int J Environ Res Public Health* 2020 Apr 23;17(8):2906 [FREE Full text] [doi: [10.3390/ijerph17082906](https://doi.org/10.3390/ijerph17082906)] [Medline: [32340107](https://pubmed.ncbi.nlm.nih.gov/32340107/)]
41. Crawford A, Serhal E. Digital Health Equity and COVID-19: The Innovation Curve Cannot Reinforce the Social Gradient of Health. *J Med Internet Res* 2020 Jun 02;22(6):e19361 [FREE Full text] [doi: [10.2196/19361](https://doi.org/10.2196/19361)] [Medline: [32452816](https://pubmed.ncbi.nlm.nih.gov/32452816/)]
42. Fagherazzi G, Goetzing C, Rashid MA, Aguayo GA, Huiart L. Digital Health Strategies to Fight COVID-19 Worldwide: Challenges, Recommendations, and a Call for Papers. *J Med Internet Res* 2020 Jun 16;22(6):e19284 [FREE Full text] [doi: [10.2196/19284](https://doi.org/10.2196/19284)] [Medline: [32501804](https://pubmed.ncbi.nlm.nih.gov/32501804/)]
43. Hong Z, Li N, Li D, Li J, Li B, Xiong W, et al. Telemedicine During the COVID-19 Pandemic: Experiences From Western China. *J Med Internet Res* 2020 May 08;22(5):e19577 [FREE Full text] [doi: [10.2196/19577](https://doi.org/10.2196/19577)] [Medline: [32349962](https://pubmed.ncbi.nlm.nih.gov/32349962/)]
44. Horowitz RI, Freeman PR. Three novel prevention, diagnostic, and treatment options for COVID-19 urgently necessitating controlled randomized trials. *Med Hypotheses* 2020 Oct;143:109851 [FREE Full text] [doi: [10.1016/j.mehy.2020.109851](https://doi.org/10.1016/j.mehy.2020.109851)] [Medline: [32534175](https://pubmed.ncbi.nlm.nih.gov/32534175/)]
45. Kannappallil T, Ma J. Digital Translucence: Adapting Telemedicine Delivery Post-COVID-19. *Telemed J E Health* 2020 Sep 01;26(9):1120-1122. [doi: [10.1089/tmj.2020.0158](https://doi.org/10.1089/tmj.2020.0158)] [Medline: [32427529](https://pubmed.ncbi.nlm.nih.gov/32427529/)]
46. Keshvardoost S, Bahaadinbeigy K, Fatehi F. Role of Telehealth in the Management of COVID-19: Lessons Learned from Previous SARS, MERS, and Ebola Outbreaks. *Telemed J E Health* 2020 Jul 23;26(7):850-852. [doi: [10.1089/tmj.2020.0105](https://doi.org/10.1089/tmj.2020.0105)] [Medline: [32329659](https://pubmed.ncbi.nlm.nih.gov/32329659/)]
47. Massaroni C, Nicolò A, Schena E, Sacchetti M. Remote Respiratory Monitoring in the Time of COVID-19. *Front Physiol* 2020 May 29;11:635 [FREE Full text] [doi: [10.3389/fphys.2020.00635](https://doi.org/10.3389/fphys.2020.00635)] [Medline: [32574240](https://pubmed.ncbi.nlm.nih.gov/32574240/)]

48. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile Health Apps on COVID-19 Launched in the Early Days of the Pandemic: Content Analysis and Review. *JMIR mHealth uHealth* 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: [10.2196/19796](https://doi.org/10.2196/19796)] [Medline: [32609622](https://pubmed.ncbi.nlm.nih.gov/32609622/)]
49. Sharma T, Bashir M. Use of apps in the COVID-19 response and the loss of privacy protection. *Nat Med* 2020 Aug 26;26(8):1165-1167. [doi: [10.1038/s41591-020-0928-y](https://doi.org/10.1038/s41591-020-0928-y)] [Medline: [32457443](https://pubmed.ncbi.nlm.nih.gov/32457443/)]
50. Thulesius H. Increased importance of digital medicine and eHealth during the Covid-19 pandemic. *Scand J Prim Health Care* 2020 Jun 02;38(2):105-106. [doi: [10.1080/02813432.2020.1770466](https://doi.org/10.1080/02813432.2020.1770466)] [Medline: [32484725](https://pubmed.ncbi.nlm.nih.gov/32484725/)]
51. Watson AR, Wah R, Thamman R. The Value of Remote Monitoring for the COVID-19 Pandemic. *Telemed J E Health* 2020 Sep 01;26(9):1110-1112. [doi: [10.1089/tmj.2020.0134](https://doi.org/10.1089/tmj.2020.0134)] [Medline: [32384251](https://pubmed.ncbi.nlm.nih.gov/32384251/)]
52. Ye J. The Role of Health Technology and Informatics in a Global Public Health Emergency: Practices and Implications From the COVID-19 Pandemic. *JMIR Med Inform* 2020 Jul 14;8(7):e19866 [FREE Full text] [doi: [10.2196/19866](https://doi.org/10.2196/19866)] [Medline: [32568725](https://pubmed.ncbi.nlm.nih.gov/32568725/)]
53. Cohen JK. Remote monitoring helps keep low-acuity COVID-19 patients at home. *Modern Healthcare*. URL: <https://www.pressreader.com/usa/modern-healthcare/20200420/281840055803238> [accessed 2021-04-11]
54. Edelman L, McConnell E, Kennerly S, Alderden J, Horn S, Yap T. Mitigating the Effects of a Pandemic: Facilitating Improved Nursing Home Care Delivery Through Technology. *JMIR Aging* 2020 May 26;3(1):e20110 [FREE Full text] [doi: [10.2196/20110](https://doi.org/10.2196/20110)] [Medline: [32412909](https://pubmed.ncbi.nlm.nih.gov/32412909/)]
55. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. *BMJ* 2020 Mar 12;368:m998. [doi: [10.1136/bmj.m998](https://doi.org/10.1136/bmj.m998)] [Medline: [32165352](https://pubmed.ncbi.nlm.nih.gov/32165352/)]
56. John O. Video consultations for triage of patients with covid-19. *BMJ* 2020 Apr 23;369:m1583. [doi: [10.1136/bmj.m1583](https://doi.org/10.1136/bmj.m1583)] [Medline: [32327408](https://pubmed.ncbi.nlm.nih.gov/32327408/)]
57. Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. *BMJ* 2020 Mar 25;368:m1182. [doi: [10.1136/bmj.m1182](https://doi.org/10.1136/bmj.m1182)] [Medline: [32213507](https://pubmed.ncbi.nlm.nih.gov/32213507/)]
58. Japan ECMONet for COVID-19. Japan ECMONet for COVID-19: telephone consultations for cases with severe respiratory failure caused by COVID-19. *J Intensive Care* 2020 Apr 7;8(1):24 [FREE Full text] [doi: [10.1186/s40560-020-00440-9](https://doi.org/10.1186/s40560-020-00440-9)] [Medline: [32288941](https://pubmed.ncbi.nlm.nih.gov/32288941/)]
59. Barsom EZ, Feenstra TM, Bemelman WA, Bonjer JH, Schijven MP. Coping with COVID-19: scaling up virtual care to standard practice. *Nat Med* 2020 May 14;26(5):632-634. [doi: [10.1038/s41591-020-0845-0](https://doi.org/10.1038/s41591-020-0845-0)] [Medline: [32405054](https://pubmed.ncbi.nlm.nih.gov/32405054/)]
60. Fitz Maurice M, di Tommaso F, Baranchuk A. Efecto de una intervención en las redes sociales en tiempos del COVID-19. *Medicina*. URL: [https://www.medicinabuenosaires.com/indices-de-2020/volumen-80-ano-2020-no-3-indice/carta\\_efecto/](https://www.medicinabuenosaires.com/indices-de-2020/volumen-80-ano-2020-no-3-indice/carta_efecto/) [accessed 2021-04-11]
61. Giansanti D, Aprile I. Is the COVID-19 Pandemic an Opportunity to Enlarge the Telemedicine Boundaries? *Telemed J E Health* 2020 Sep 01;26(9):1123-1125. [doi: [10.1089/tmj.2020.0159](https://doi.org/10.1089/tmj.2020.0159)] [Medline: [32456559](https://pubmed.ncbi.nlm.nih.gov/32456559/)]
62. Hau YS, Kim JK, Hur J, Chang MC. How about actively using telemedicine during the COVID-19 pandemic? *J Med Syst* 2020 Apr 30;44(6):108 [FREE Full text] [doi: [10.1007/s10916-020-01580-z](https://doi.org/10.1007/s10916-020-01580-z)] [Medline: [32350626](https://pubmed.ncbi.nlm.nih.gov/32350626/)]
63. Jamil B. Clinical features, diagnosis and management of COVID-19 patients in the outdoor setting. *J Pak Med Assoc* 2020 May;70(Suppl 3)(5):S52-S55. [doi: [10.5455/JPMA.12](https://doi.org/10.5455/JPMA.12)] [Medline: [32515378](https://pubmed.ncbi.nlm.nih.gov/32515378/)]
64. Nair SC, Satish KP, Sreedharan J, Muttappallymyalil J, Ibrahim H. Improving Health Literacy Critical to Optimize Global Telemedicine During COVID-19. *Telemed J E Health* 2020 Nov 01;26(11):1325-1325. [doi: [10.1089/tmj.2020.0175](https://doi.org/10.1089/tmj.2020.0175)] [Medline: [32463318](https://pubmed.ncbi.nlm.nih.gov/32463318/)]
65. Trethewey SP, Beck KJ, Symonds RF. Video consultations in UK primary care in response to the COVID-19 pandemic. *Br J Gen Pract* 2020 May;70(694):228-229. [doi: [10.3399/bjgp20X709505](https://doi.org/10.3399/bjgp20X709505)] [Medline: [32354807](https://pubmed.ncbi.nlm.nih.gov/32354807/)]
66. Wei Y, Li G, Li J, Chu J, Yue H, Zhang F, et al. Outpatient strategies and the changing SARS-CoV-2 epidemic. *ERJ Open Res* 2020 Apr 17;6(2):00111-2020 [FREE Full text] [doi: [10.1183/23120541.00111-2020](https://doi.org/10.1183/23120541.00111-2020)] [Medline: [32440516](https://pubmed.ncbi.nlm.nih.gov/32440516/)]
67. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: Evidence from the field. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
68. Hunter A, Facey K, Thomas V, Haerry D, Warner K, Klingmann I, et al. EUPATI Guidance for Patient Involvement in Medicines Research and Development: Health Technology Assessment. *Front Med (Lausanne)* 2018 Sep 6;5:231 [FREE Full text] [doi: [10.3389/fmed.2018.00231](https://doi.org/10.3389/fmed.2018.00231)] [Medline: [30238004](https://pubmed.ncbi.nlm.nih.gov/30238004/)]
69. Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. *NPJ Digit Med* 2018;1:20172 [FREE Full text] [doi: [10.1038/s41746-017-0002-4](https://doi.org/10.1038/s41746-017-0002-4)] [Medline: [31304346](https://pubmed.ncbi.nlm.nih.gov/31304346/)]
70. Su D, Michaud TL, Estabrooks P, Schwab RJ, Eiland LA, Hansen G, et al. Diabetes Management Through Remote Patient Monitoring: The Importance of Patient Activation and Engagement with the Technology. *Telemed J E Health* 2019 Oct 01;25(10):952-959. [doi: [10.1089/tmj.2018.0205](https://doi.org/10.1089/tmj.2018.0205)] [Medline: [30372366](https://pubmed.ncbi.nlm.nih.gov/30372366/)]
71. Jain VK, Iyengar K, Vaish A, Vaishya R. Differential mortality in COVID-19 patients from India and western countries. *Diabetes Metab Syndr* 2020 Sep;14(5):1037-1041 [FREE Full text] [doi: [10.1016/j.dsx.2020.06.067](https://doi.org/10.1016/j.dsx.2020.06.067)] [Medline: [32640415](https://pubmed.ncbi.nlm.nih.gov/32640415/)]

72. Moore JT, Ricaldi JN, Rose CE, Fuld J, Parise M, Kang GJ, COVID-19 State, Tribal, Local, and Territorial Response Team. Disparities in Incidence of COVID-19 Among Underrepresented Racial/Ethnic Groups in Counties Identified as Hotspots During June 5-18, 2020 - 22 States, February-June 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 21;69(33):1122-1126 [FREE Full text] [doi: [10.15585/mmwr.mm6933e1](https://doi.org/10.15585/mmwr.mm6933e1)] [Medline: [32817602](https://pubmed.ncbi.nlm.nih.gov/32817602/)]
73. Muñoz-Price LS, Nattinger AB, Rivera F, Hanson R, Gmehlin CG, Perez A, et al. Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Netw Open* 2020 Sep 01;3(9):e2021892 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.21892](https://doi.org/10.1001/jamanetworkopen.2020.21892)] [Medline: [32975575](https://pubmed.ncbi.nlm.nih.gov/32975575/)]

## Abbreviations

**RMT:** remote monitoring technology

*Edited by L Buis; submitted 02.10.20; peer-reviewed by Y Tian, N Benda, S Mukherjee; comments to author 30.11.20; revised version received 03.03.21; accepted 22.03.21; published 20.04.21.*

*Please cite as:*

*Houlding E, Mate KKV, Engler K, Ortiz-Paredes D, Pomey MP, Cox J, Hijal T, Lebouché B*

*Barriers to Use of Remote Monitoring Technologies Used to Support Patients With COVID-19: Rapid Review*

*JMIR Mhealth Uhealth* 2021;9(4):e24743

URL: <https://mhealth.jmir.org/2021/4/e24743>

doi: [10.2196/24743](https://doi.org/10.2196/24743)

PMID: [33769943](https://pubmed.ncbi.nlm.nih.gov/33769943/)

©Elizabeth Houlding, Kedar K V Mate, Kim Engler, David Ortiz-Paredes, Marie-Pascale Pomey, Joseph Cox, Tarek Hijal, Bertrand Lebouché. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 20.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# A Smartphone App Designed to Empower Patients to Contribute Toward Safer Surgical Care: Qualitative Evaluation of Diverse Public and Patient Perceptions Using Focus Groups

Stephanie Russ<sup>1</sup>, PhD; Nick Sevdalis<sup>1</sup>, PhD; Josephine Ocloo<sup>1</sup>, PhD

Centre for Implementation Science, King's College London, London, United Kingdom

**Corresponding Author:**

Stephanie Russ, PhD

Centre for Implementation Science

King's College London

De Crespigny Park

London, SE58AF

United Kingdom

Phone: 44 2078480683

Email: [stephanie.russ@kcl.ac.uk](mailto:stephanie.russ@kcl.ac.uk)

## Abstract

**Background:** MySurgery is a smartphone app designed to empower patients and their caregivers to contribute toward safer surgical care by following practical advice to help reduce susceptibility to errors and complications.

**Objective:** The aim of this study is to evaluate service users' perceptions of MySurgery, including its perceived acceptability, the potential barriers and facilitators to accessing and using its content, and ideas about how to facilitate its effective implementation. The secondary aim is to analyze how the intended use of the app might differ for diverse patients, including seldom-heard groups.

**Methods:** We implemented a diversity approach to recruit participants from a range of backgrounds with previous experience of surgery. We aimed to achieve representation from seldom-heard groups, including those from a Black, Asian, and minority ethnic (BAME) background; those with a disability; and those from the lesbian, gay, bisexual, transgender, queer (LGBT+) community. A total of 3 focus groups were conducted across a 2-month period, during which a semistructured protocol was followed to elicit a rich discussion around the app. The focus groups were audio recorded, and thematic analysis was carried out.

**Results:** In total, 22 individuals participated in the focus groups. A total of 50% (n=11) of the participants were from a BAME background, 59% (n=13) had a disability, and 36% (n=8) were from the LGBT+ community. There was a strong degree of support for the MySurgery app. The majority of participants agreed that it was acceptable and appropriate in terms of content and usability, and that it would help to educate patients about how to become involved in improving safety. The *checklist-like* format was popular. There was rich discussion around the accessibility and inclusivity of MySurgery. Specific user groups were identified who might face barriers in accessing the app or acting on its advice, such as those with visual impairments or learning difficulties and those who preferred to take a more passive role (eg, some individuals because of their cultural background or personality type). The app could be improved by signposting further specialty-specific information and incorporating a calendar and notes section. With regard to implementation, it was agreed that use of the app should be signposted before the preoperative appointment and that training and education should be provided for clinicians to increase awareness and buy-in. Communication about the app should clarify its scientific basis in plain English and should stress that its use is optional.

**Conclusions:** MySurgery was endorsed as a powerful tool for enhancing patient empowerment and facilitating the direct involvement of patients and their caregivers in maintaining patient safety. The diversity approach allowed for a better understanding of the needs of different population groups and highlighted opportunities for increasing accessibility and involvement in the app.

(*JMIR Mhealth Uhealth* 2021;9(4):e24065) doi:[10.2196/24065](https://doi.org/10.2196/24065)

**KEYWORDS**

patient safety; mobile health; patient involvement; perioperative care; smartphone app; mobile phone



## Introduction

### Background

Health care is being called to better embrace the potential of digital technology for transforming patient care. The unprecedented spread of mobile digital technologies and their potential to address health priorities has evolved into the field of mobile health, which can be defined as “medical and public health practice supported by mobile devices” [1]. Smartphone apps, as an example of such technology, have emerged as a key device for communicating health information at scale and for improving patient empowerment, which is an important long-term objective of health care internationally [2-6]. A growing evidence base suggests that smartphone technology may play an important role in improving mental and physical health outcomes for a range of patient groups in both high- and low-income countries [7-11]. In the future, the use of smartphone technology to enable patients to be more involved in their care and to connect them with their providers outside the clinic is likely to be even more important. The recent COVID-19 pandemic offers a prime example of a health system response that has relied on the implementation of remote and largely digital solutions for delivering and receiving care.

This paper focuses on MySurgery, an app for use within the context of surgical care. MySurgery is a smartphone app designed to empower patients and their caregivers to play a role in improving safety in surgical care in the National Health Service (NHS) of the United Kingdom. MySurgery was created by a multidisciplinary team of clinicians, patient safety experts, and patient and public representatives and is available for free download on Apple devices from the App Store; to date, it has had more than 6000 downloads [12]. The animated, jargon-free app is centered around key areas of evidence-based surgical risk, such as medication management, consent and identification, wound care, falls, and hand hygiene (Figure 1). The app is designed to be used as an optional supplementary tool, rather than replacing any existing surgical educational material or safety procedures. It provides practical step-by-step advice on the actions patients and their caregivers can take to help mitigate risks, including warning signs to look out for, things to do, information to provide, and questions to ask. The objective is to encourage behaviors that will help to reduce a patient’s

susceptibility to error, for example, by flagging up inconsistencies or omissions, by encouraging behaviors that reduce the risk of infection or falls, and by providing information to allow better clinical decision making. It also helps patients to prepare optimally before and after surgery. Previous research has identified these as the kind of *safety-related behaviors* that should be incorporated into interventions designed to enhance patient involvement in safety [13-15]. The interventional rationale behind MySurgery is much in keeping with that of enhanced recovery programs, which are evidence-based perioperative programs that employ a multidisciplinary and multimodal approach based on implementing a checklist of actions that help resolve issues that delay recovery and cause complications [16]. These programs have been shown to significantly reduce morbidity, mortality, and length of hospital stay [17,18]. In a recent pilot evaluation of MySurgery with a diverse group of 42 surgical patients, the app received positive feedback [19].

The potential benefits of mobile apps such as MySurgery can only be realized if they are deemed acceptable by the end user and the initial intent to use the intervention translates to actual user engagement with it. Engagement with a digital intervention can be conceptualized as the extent (eg, amount, frequency, duration, and depth) of use and a subjective experience characterized by attention, interest, and affect [20]. Understanding the potential barriers to and facilitators of engagement before the implementation of an intervention is critical for successful implementation. A number of technology adoption models have been proposed that define the kind of factors that will influence perceived acceptability and intention to use digital interventions, ranging from perceived ease of use, social influences and subjective norms, requirement for access to personal data and related security concerns, perceived usefulness, and perceived behavioral control (or perceived ability to carry out a given behavior promoted by the app) [21-23]. The influence these factors have on engagement will vary based on different features of the app at hand (eg, how much personal data input an app requires); however, it is also likely to vary for different patient groups. For example, those from different cultures or with certain disabilities may experience different barriers to and facilitators of using the app. These potential group differences have rarely been explored in app usability and evaluation studies.

Figure 1. MySurgery app screenshots.



## Objectives

This study has 2 aims. The first aim is to explore in-depth views about MySurgery held by individuals who had previously undergone surgery, using focus groups. Focus groups are a fruitful way to gather the perceptions and experiences of groups of service users that offer them more control of the interaction [24]. Specifically, we aim to understand the perceptions related to the following:

1. The content and usability of MySurgery, including its perceived acceptability and usefulness
2. The perceived potential impact of using MySurgery on the doctor-patient relationship
3. Potential barriers to and facilitators of using MySurgery and associated strategies for overcoming any barriers
4. Potential strategies for introducing MySurgery into surgical care pathways in the UK NHS

The second aim is to implement a diversity and inclusion approach for recruitment, which would allow the research to address weaknesses in the literature to date. There are long-standing criticisms about tokenism and the lack of diversity in involving patients and the public in health care and health care research [25,26]. By targeting seldom-heard groups in the recruitment process, we aim to ensure that their representation will enable us to gather feedback from a broader spectrum of the population for whom the app is designed. We hypothesized that the views held regarding the app, specifically perceived barriers to using it in practice, would vary according to sociocultural factors and disability status. Therefore, we felt that a diversity approach was important.

## Methods

### Participants

To implement the diversity and inclusion approach, as described later, the recruitment of individuals for the focus groups took place via 2 routes. The first route was to advertise on *People in Research*, which is a UK-hosted website designed to provide opportunities for the public to become involved in NHS, public health, and social care research. A bespoke study advert (Multimedia Appendix 1) was published on the website, offering participation in the focus groups or involvement in the project steering group (the project steering group met biannually to oversee the wider research program and consisted of public and patient representatives, clinicians, and patient safety scientists. It is not described further here). Interested individuals could respond directly to the study team to express their interest in taking part. The second route was to email the project advert to individuals who had asked to be informed of upcoming research opportunities through the Patient and Public Involvement (PPI) Group at the Centre for Implementation Science (CIS), King's College London, where the research was hosted. The criteria stipulated that individuals should only apply if they were over the age of 18 years and could speak, understand, and read English. We also set criteria that they should have experienced surgery within the past 5 years, such that they could draw on their experience to respond to the questions, which were focused on patient involvement in surgical care.

### Diversity and Inclusion Approach

As part of the study aims, we implemented a diversity and inclusion approach whereby we sought to include representation from individuals from seldom-heard groups. This approach included the following elements:

1. We used the Equality Act 2010 [27] to inform our conceptualization of diversity, which sets out 9 protected characteristics that may be the subject of discrimination. We chose to focus on 3 of these characteristics, which relate to groups that have traditionally been underrepresented in health care research: disability, race, and sexual orientation [28,29].
2. Using the study advert (Multimedia Appendix 1), we invited anyone with previous experience of surgery to participate in the project. However, we also set out *desired criteria* focused on the aforementioned 3 protected characteristics, encouraging the following groups in particular to apply:

- those with a disability; those from a Black, Asian, and minority ethnic (BAME) group; and those from the lesbian, gay, bisexual, transgender, and queer (LGBT+) community.
3. The CIS PPI Group (one of the aforementioned routes through which we recruited participants) had just been set up using a local community-based diversity approach, which had drawn upon the Equality Act. Therefore, it had a contacts list with a wide diversity of groups, which we felt would assist in the recruitment of a diverse sample.
4. We drew upon the cultural competence and expertise of JO as a researcher from a BAME background, with long-standing expertise working on diversity issues.
5. An *equality monitoring form* (Multimedia Appendix 2) was used to monitor the diversity of responses to our advert with sections of the form based on categories in the Equality Act 2010. We then aimed to ensure that we represented a diverse range of groups by selecting at least 50% of our sample from those stipulating that they held 1 of the 3 protected characteristics.
6. Before the commencement of the focus groups, the *rules for engagement* were discussed and agreed upon by all, with the aim of breaking down any potential barriers to inclusion. These included practical points such as turning off mobile phones and points on communication style, such as not speaking over one another, allowing everyone to speak, not needing permission to speak, and being able to respond directly to one another. Participants were reassured that there were no right or wrong answers and that we were not trying to reach a consensus but rather to gather information and understand a range of different viewpoints.
7. To encourage participation from diverse groups, all participants received a one-off payment of UK £40 (US \$56), a lunch following the focus group, and their travel expenses.

### Design

Three 90-minute focus groups were conducted across a 3-month period (November 11, 2017, to January 1, 2018). Each focus group had a maximum capacity of 10 participants and was facilitated by 2 researchers with previous experience and training in facilitating focus groups (SR and JO). The project was reviewed and approved by the King's College London Ethics Committee (REF: MR/17/18-108).

### Materials

#### *Equality Monitoring Team*

To measure the diversity of our sample, we asked participants to complete a standardized equality monitoring form (Multimedia Appendix 2), which captured demographic information, including age, ethnicity, disability status, and sexuality.

#### *Bespoke Focus Group Questionnaire*

To better understand our sample, we asked individuals to complete a short bespoke questionnaire (Multimedia Appendix 3) that captured information relating to their previous experience of having had surgery (including date of last surgery and how many procedures they had had in total) and their familiarity with using smartphone apps.

### ***Semistructured Discussion Guide***

A semistructured discussion guide was devised to assist the facilitators in guiding the conversation during the focus groups. It included questions and prompts structured around the study aims and a rough schedule to ensure everything was covered in the available time, building in room for refreshments and comfort breaks ([Multimedia Appendix 4](#)). Thinking about taking an inclusive approach was particularly important; this meant thinking about room access, options for any special diets, developing ground rules to ensure everyone was able to participate equally, and building specific questions to prompt discussion about how different groups might be able to use the app. For example: Will the app be acceptable to all?; For whom will the app be less useful?; What about people who do not own or use a smartphone or iPad?; How can we get around the issues identified?; Can you think of alternative ways of delivering this information?; and Should this impact on the decision to take the app forward?

### **Procedure**

Individuals who responded to the study advert were sent a more detailed study information sheet in the post or via email. Those who were still interested in being involved were sent the study questionnaires (equality monitoring form and bespoke focus group questionnaire) and asked to bring the completed forms to their focus group, which was booked for 1 of 3 dates in November or December 2017. If they had access to an Apple device, they were asked to download and familiarize themselves with the MySurgery app before attending their focus group.

On the day of the focus group, participants arrived 30 minutes before the commencement of the discussion to allow time to complete their informed consent, meet the other group members, and get refreshments. They could also use this time to familiarize themselves with the MySurgery app (on study iPads), if they had not been able to at home or if they wanted to refresh their memory. The focus group itself was preceded by a 10-minute presentation by the 2 facilitators, to introduce themselves and the aims of the study and to provide a brief background on the development and objectives of the MySurgery app. Following this, the participants introduced themselves, and the *rules for engagement* in the focus group discussion were outlined. Relevant ethical issues such as how data would be stored and used and the participants' right to withdraw were also reviewed.

The discussions were audio recorded (with participants' permission) for subsequent transcription for analysis purposes

and lasted for 90 minutes, with a 10-minute comfort and refreshment break.

### **Data Analysis**

The audio recordings of the 3 focus groups were transcribed to allow a qualitative thematic analysis of the discussion. This was completed separately by 2 researchers (SR and JO) with expertise in qualitative data analysis. An inductive approach to the analysis was undertaken whereby the raw textual data were used to extract themes based on exploratory interpretation of the material, providing a summary of its content. This approach enables the development of a theory about the perceptions held by the focus group participants, with clear links to the research objectives [30]. Practically, this was achieved through the following steps:

1. Focus group transcripts were combined and rearranged to group answers together for each interview protocol question.
2. For each question, we noted the main ideas that were raised in the discussion.
3. We reviewed the main ideas (across all questions) to identify ideas that were raised again and again or engendered a particularly rich discussion.
4. The researchers performed critical thinking about these recurring ideas to identify emergent themes.
5. We identified quotations from the transcripts to illustrate each theme.

To establish coding agreement between the researchers, initially half of the transcripts were analyzed by both researchers, and the extracted themes were compared and agreed upon. Subsequently, the remaining half of the transcripts were divided equally, and the researchers came together at the end to agree on the final emergent themes.

## **Results**

### **Participants**

A total of 22 individuals participated in the focus groups (group 1: n=8, 36%; group 2: n=9, 41%; and group 3: n=5, 23%). The participant characteristics are summarized in [Table 1](#). There was an even spread of males and females, and the groups were diverse according to age, ethnicity, disability, and caregiver status. Just over one-third (n=8, 36%) of the sample was from the LGBT+ community. All participants had undergone surgery within the past 5 years, excluding one individual who was awaiting upcoming surgery.

**Table 1.** Participant characteristics (N=22).

Characteristic	Participants, n (%)
<b>Sex</b>	
Male	10 (45)
Female	12 (54)
<b>Age (years)</b>	
18-24	1 (5)
25-34	2 (9)
35-44	2 (9)
45-54	4 (19)
55-65	9 (43)
>65	3 (14)
<b>Ethnicity</b>	
White	11 (50)
<b>BAME<sup>a</sup></b>	11 (50)
Asian/Asian British	3 (14)
Black/African/Caribbean/Black British	5 (23)
Mixed/multiple ethnic groups	3 (14)
<b>Sexual orientation</b>	
Heterosexual	14 (64)
Gay	3 (14)
Lesbian	1 (4)
Other	4 (18)
<b>Disability</b>	
Yes	13 (59)
No	9 (41)
<b>Caregivers</b>	
Yes	9 (41)
No	13 (59)
<b>Number of previous surgeries</b>	
0	1 (4.5)
1-2	9 (41)
3-4	10 (45)
≥5	2 (9)
<b>Smartphone user</b>	
Yes	20 (91)
No	2 (9)
<b>Would you use a smartphone for health-related purposes</b>	
Yes	19 (86)
No	3 (14)

<sup>a</sup>BAME: Black, Asian, and minority ethnic.

## Thematic Analysis of Focus Group Discussions

The following themes were extracted from the focus group transcripts. Illustrative quotes are provided in [Multimedia Appendix 5](#).

### Perceptions Regarding Patient Involvement in Safety

There was a strong sense of support for the concept that patients should be closely involved in the safety of their own care. Some participants were very confident about this notion:

*Nothing about us, without us.*

*It's about taking responsibility for your care and we need to be moving in that direction.*

A number of areas were highlighted where most patients do naturally become involved in the safety of their care, including complying with instructions around food, drink, and medications; attending preoperative appointments; and asking questions. However, it was suggested that there are areas where it might be difficult to become involved in safety, particularly where it involves conflict behaviors such as asking many questions, challenging the decisions of health care workers, or highlighting inconsistencies, suggesting that this may have a negative impact on the doctor--patient relationship and erode trust. For example, in some cultures, it is frowned upon to question a figure of authority (it was suggested that some minority groups might feel that such behaviors might affect the standard of care they receive), those with less assertive personality types might simply feel unable to engage in these behaviors, and others may feel more comfortable with the passivity of the traditional patient role:

*What if you've got a patient who actively just wants to be told what to do, doesn't want to have that responsibility.*

It was agreed that for these areas particularly, facilitation to become involved in safety would be needed, including encouragement and empowerment from the health care workers themselves, information around how to contribute, and education around the broader safety-related areas patients and their advocates can feasibly influence:

*Patients can get involved in patient safety but they need encouragement by clinicians and to be told in what ways they can contribute.*

The importance of the involvement of family, friends, and caregivers was also highlighted, particularly for individuals who were too unwell to be involved themselves.

### MySurgery App: Concept, Content, and Usability

MySurgery was endorsed as a positive step toward patient empowerment and an acceptable approach to facilitating patient involvement in safety. In a practical sense, the app was recognized to help participants prepare before and after surgery, for example, by detailing what to take into hospital and how to care for surgical wounds. From a safety perspective, participants agreed that by enhancing their knowledge of surgical risks and listing specific actions for mitigating them, the app shows the possibility of becoming more involved in safety and very real

opportunities for avoiding errors or complications, regardless of the type of procedure being performed:

*I think they (patients) ought to know what they can and can't do, and if they can ask questions and it's a two-way communication. That's what this app seems to do, it shows people the possibilities, shows patients the possibilities.*

It was agreed that MySurgery raises safety matters that might be of relevance but that might not have been thought about and that patients could address these issues using their own approach and style. In particular, MySurgery was recognized as a useful communication tool, informing patients about the areas they may wish to discuss with their health care team and about the information they should provide:

*It just empowers you in a way because it doesn't mean you have the confidence to say, have you washed your hands, but maybe next time you'll be able to or maybe you can attempt to, or know that that's important, or discuss it in a different way with the health worker.*

It was stressed, however, that unless users of the app understood the context of patient empowerment and involvement within which the app sits, they might not fully grasp its relevance and importance, which may prevent buy-ins.

In terms of detail, participants liked the level of content in MySurgery in that it was *not information overload* and was clean, clear, and simple, with no jargon, even for individuals who struggle with medical information. The *checklist-like* layout of the app content was approved, as participants liked the feeling that they could cross items off their list and that they had *covered off all of the important things*. As a quick aide-memoire, the *top 10 things to remember* tab within the app was also popular. The app content was deemed generic enough to be appropriate for all surgical patients; however, it was discussed that those having very minor procedures or multiple procedures may only need to check the *top 10* section or may only be interested in 1 or 2 sections. In terms of usability, the app was described generally as being user friendly, easy to navigate, and nicely animated:

*I liked it, I thought it was very user friendly, it wasn't information overload so when you opened it and thought, oh god, I'm going to have to sit and spend ages reading through loads of...but it's quite interactive and it's short.*

Some suggestions were made to improve MySurgery. Currently, the app requires you to work through it in a set order, meaning that later sections cannot be accessed until earlier sections have been completed. This was not a popular feature, as participants felt that they wanted to access whichever section looked most relevant to them, skipping other sections (eg, those having repeat or very minor procedures may only want to look at 1 or 2 sections). It was widely agreed that this limitation should be addressed. Others suggested that although the level of content was good, some users would like more details and information, perhaps specific to the procedure they are having done, and links and signposts to this kind of information could be built into the app:

*How about links to information about the procedure you are having, or somewhere to find more information?*

Some further suggestions were made around personalizing the app and incorporating more general support for managing one's care, including the addition of a calendar to provide alerts for

appointments or medications, a contacts list to record names of key clinicians, and an area in which to make notes. Finally, it was suggested that if you could enter your surgery date into the app, alerts could then be provided to remind you to check it. The suggested improvements to MySurgery are summarized in [Textbox 1](#).

**Textbox 1.** Summary of findings: recommendations for the development and implementation of MySurgery.

<p>Content</p> <ul style="list-style-type: none"> <li>• Include links and signposts to further resources and specialty-specific content.</li> <li>• Include a calendar for inputting appointments with an integrated alert system, for example, “check you have everything before attending hospital tomorrow.”</li> <li>• Include an area to make notes, for example, to record clinicians' names and important phone numbers.</li> <li>• App store information—make it clear that the use of the app is optional and noninterventional (ie, education only) and communicate the objectives of the app in plain English.</li> </ul> <p>Usability</p> <ul style="list-style-type: none"> <li>• Remove the requirement to work through the app in a set order.</li> </ul> <p>Accessibility</p> <ul style="list-style-type: none"> <li>• Develop audio and easy-to-read versions of the app.</li> <li>• Make MySurgery available for Android devices.</li> <li>• Distil the information into appropriate format for hard copies of the app, for example, booklet, leaflet, or posters.</li> <li>• Make MySurgery available in different languages.</li> <li>• Include information around where to find help in using the app.</li> </ul> <p>Implementation strategy</p> <ul style="list-style-type: none"> <li>• MySurgery should be recommended to patients before their preoperative appointment.</li> <li>• Promote the use of the app through inclusion in the preop letter, via posters in primary care General Practice surgeries and pharmacies, and by displaying on the television screens in hospital wards.</li> <li>• Educate clinicians to secure buy-in and promote the use of the app.</li> <li>• Involve clinical teams and patients in adapting the app to specific surgical specialties or hospital units.</li> <li>• Include information around how to download the app.</li> <li>• Include guidance on how to use the app in practice.</li> </ul>
---

### Accessibility

The degree to which MySurgery is accessible and inclusive was an important point of discussion. If a safety intervention is accessible to some but not others, there is a moral issue to be considered in terms of equity, with some being unable to access information that might enhance their care. The viewpoints around this were varied, and the discussion was rich. MySurgery is currently available on Apple devices, meaning that those who cannot access an Apple device will not be able to benefit from the information within the app. Examples of groups falling into this category are those with Android devices, those who do not have access to a smartphone or tablet, those who find apps and technology difficult to use, those who do not speak English, those with visual impairments, those with learning difficulties or dementia, and those within institutions:

*I think that the app is very clear and the language is very clear however there might be groups of society*

*who have a lot of surgery that find it difficult. I'm talking about learning disabilities, Alzheimer's and so on..., or people that have had neurological trauma and also people who might already be very unwell in hospital.*

It is very important to consider how the accessibility of technology to these diverse groups can be enabled or improved. For example, MySurgery will soon be available on Android devices. Other adaptations might include developing an audio version of the app, developing easy-to-read versions, translating the app into different languages (so far, the app is available in English and French), developing paper-based versions, and considering whether the app needs to be adapted for situations where it will be used by a caregiver as opposed to the patient themselves (eg, when the patient is a child, has dementia, or is too sick to use the technology themselves).

The use of MySurgery by older individuals (ie, 65+ years) was also discussed, as there was a question around how accessible

smartphone technology is to this population. There was widespread agreement that although certain individuals in this bracket may not be as familiar with smartphone technology as others, older individuals in general are increasingly comfortable with technology and are equally likely to own and use a smartphone or tablet (this is supported by the national survey data) [31]. Thus, it should not be assumed that older individuals will struggle to access the intervention; however, alternative formats, such as a hard copy of the app content, might be preferred:

*Don't make assumptions about the elderly...[be]cause there is a presumption that the elderly don't like technology. I've spoken to some people in their 70s, I never say to them, do you know how to use an iPad, you just hand it over.*

Another group of individuals were identified who have access to technology but would not wish to engage with this kind of information or do not want to know about risk—the traditionally passive patient described earlier. This reinforces the point that the use of technology should not be forced but rather presented as an option for enhancing care.

As a positive endorsement of MySurgery, there was agreement among most members of the group that the identified barriers to accessibility, although important to address, should not prevent investment in or implementation of the app, particularly where substantial improvements in the efficiency and safety of care might be realized by its use for so many:

*You cannot with one thing reach all people. You probably won't be able to reach a certain percentage, but the point is it's actually much more efficient for most.*

The group agreed, therefore, that as an intervention, MySurgery should be taken forward but with the support of efforts to make adaptations to improve adoption and accessibility where necessary.

### **Implementation**

Participants agreed on some key factors that would be important to the implementation success of MySurgery (see [Textbox 1](#) for a summary of these findings). First is the adoption of a multifaceted approach that targets patients before their preoperative appointment. The preoperative appointment typically takes place within a few weeks before surgery and is an opportunity for the patient (who may attend with a relative or friend) to discuss their upcoming surgery and to ask questions. Viewing the app before the appointment would mean they would have time to digest the information and to prepare any questions or information they feel are relevant, also giving them time to take any necessary actions before their procedure. A number of suggestions for signposting patients toward the app at this time point were made. These included mentioning the app in the preoperative letter (or including a flyer with the letter) with information about it and instructions about how to download it; providing information about the app on posters or on-screen adverts in pharmacies, GP surgeries, and hospital waiting rooms; or having tablets available in waiting rooms for patients to explore the app:

*The letter that goes out to you says, you might want to look at this app first of all and if you have any worries about patient safety, bring them up in the interview, in the consultation. That is embedded in part of the consultation process and accepted.*

The app could also be made available in paper version for those who find it difficult to use the technology. It will be important to make use of the NHS digital champions linked to trusts to help support implementation and to aid with adaptations such as this that might make it more accessible.

The second factor raised by the participants was that the implementation process should educate not only patients about the app but also clinicians. Fundamentally, MySurgery, like any safety intervention, should be something that all stakeholders, including clinical staff, are aware of and bought into to optimize its effectiveness. In the case of MySurgery, having buy-in and awareness from the clinicians would mean they could recommend the app if the patient had not yet heard about it, promoting uptake and endorsing its validity, and in doing so encourage patients to engage in safety-related behaviors. This sends the signal to the patient that their involvement is considered appropriate, as opposed to feeling that they are challenging or questioning the clinician's ability. It would also help to reduce any potential strain on the patient--doctor relationship resulting from a patient mentioning an intervention the clinician is unaware of:

*There's a dual process going on here. As well as education of patients and empowering them, it's also the education of the clinicians and I think it's imperative that you speak to clinicians about it, so the onus isn't on you to have to bring it up.*

The next point around implementation focused on the content of the communications put out around the app. One potential untoward effect identified during the discussion was that the nature of the app content, which is focused on patient behaviors that can mitigate risks to safety, may be uncomfortable or even anxiety provoking for some patients. A second potential issue was that patients may feel that they *have* to use the app but may potentially struggle with accessing it or using the technology. To address these concerns, it was agreed that several points should be set out clearly in any promotional material around the app, including preoperative letters, poster or flyers, verbal communications, the App Store description of the app, and the introductory content of the app itself. It should be made clear that use of the app is supported by the NHS but that it is optional and supplementary to the safety procedures already in place. The context and scientific basis of the app (including the context of patient empowerment in which it sits) should be clarified in plain jargon-free English such that users understand why the use of the app may be beneficial and precisely what the app is trying to achieve. It should be made clear how to download the app, how to access help in using the app, and how the app content can be accessed if the users do not have a smartphone or tablet. Finally, examples of how to use the app in practice should be set out to give users ideas about where and how they might apply it to their care:



*If you want some help to access this app or would like to discuss the information, you could contact your local A, B, C...*

Finally, there was some conversation around creating specialty-specific versions of MySurgery, which are tailored to the specific risk profiles of different surgical specialties with more specific information around procedures and related recovery advice. There was also the suggestion of tailoring the app for specific hospital departments or units and integrating it with existing protocols and procedures to engender buy-in from staff and to help streamlining.

## Discussion

### Principal Findings

MySurgery is a smartphone app designed to empower patients and their caregivers to contribute toward safer surgical care. The animated checklist-type tool educates patients about simple behaviors they can undertake to mitigate a number of key evidence-based risks that are relevant to any hospital-based surgical procedure, with the objective of preventing avoidable surgical and medical complications. In this study, we conducted focus groups with 22 diverse service users who had experience of surgery to explore in-depth perceptions about the acceptability of the app, approaches to aid effective implementation, and strategies to address barriers to inclusion and accessibility.

As a general concept, participants supported the idea of patient involvement in safety but felt that education around *how* to become involved was important and that some would need assistance in doing so, which is in line with previous research [32]. MySurgery was endorsed as an acceptable approach to facilitating patients to become involved in safety by setting out key areas of risk that patients are able to influence (but which they may not have been aware of) and the safety-related behaviors they can participate in to mitigate their risk. The app was deemed acceptable in terms of content and usability, and participants agreed that by becoming better informed and understanding where problems might arise, use of the app should assist communication with health care professionals. It was also deemed probable that use of the app would reduce susceptibility to avoidable errors, given that patients would be more likely to capture errors themselves or would be more active in helping to avoid error-inducing conditions by following the advice. In this sense, patients would act as an extra safeguard, adding to the available resources in the health care system for improving safety. Some important suggestions for improvement to the app were provided, both in terms of content (eg, including links to more detailed procedure-specific information) and usability (eg, making it possible for users to access all sections of the app rather than having to work through it in a specified order).

The study also offered important insights into the diversity and inclusivity of its end users in terms of how to make the app accessible to as many patients as possible. By virtue of the recruitment approach undertaken, we achieved a diverse sample, including representation from seldom-heard groups such as those with a disability and those from a BAME background. This contributed toward a rich and varied conversation with reflections around the acceptability of MySurgery and barriers

to its use from many different perspectives. There were some examples in which the app was deemed to be well set up in terms of accessibility, such as the level of detail, animation, and avoidance of jargon, making it accessible to those who might not be comfortable or familiar with medical information. However, the discussion also identified areas where further development is required to make the app accessible to certain groups or where barriers to using the information in the app may arise.

Three key groups of individuals were identified who may experience different kinds of barriers to using MySurgery. The first were those who would like to use the app but cannot, for very practical reasons, access it in its current format, for example, those with visual impairments, those with learning difficulties, those who do not have access to a smartphone or tablet or who find technology difficult to use, those in prison or other institutions, and those who do not speak English. Various adaptations to the content and presentation of MySurgery have been suggested to address the barriers to inclusion, which should be explored and built into the ongoing development of the app. The issue of *digital exclusion* has been illuminated during the recent COVID-19 pandemic, which has seen those in certain groups, such as those who cannot afford the data required to download and interact with apps, disproportionately set back in a number of respects [33]. Alternative approaches to accessing the content of apps, such as MySurgery, must be made available to such groups.

The second group identified were those who would like to use the app but who felt they would encounter difficulties in acting on the information and advice provided. This may be due to *intrapersonal barriers* (ie, relating to the user of the app themselves), for example, if they are shy or feel psychologically vulnerable in some way, which means they find it difficult to become involved; *interpersonal barriers* (ie, between the user of the app and health care professionals), for example, where patients or professionals struggle to express themselves in a way that is easy to understand; or *cultural barriers* (*on behalf of the user of the app and/or the health care professional*), for example, where some patients from certain cultures might deem it inappropriate to question a person of authority or if health care professionals are resistant to patient involvement. The extent to which the barriers impede patient involvement may vary from behavior to behavior. For example, some behaviors recommended by the app can be completed privately with very little interaction with health care workers, for example, ensuring nonslip footwear is worn and caring appropriately for surgical wounds following discharge, whereas other behaviors rely heavily on effective communication and may be more strongly influenced by the barriers described, for example, providing information about medicines and medical history or checking if health care workers have washed their hands. Previous work on patient involvement in safety predicts that these kinds of barriers will arise and outlines the importance of overcoming them by understanding alternative approaches to empowering patients [13]. This will require ongoing work with these patients to understand which approaches to empowerment would be acceptable to them. A final group was identified who may not wish to use the app at all, for example, because they find it

anxiety provoking or because they prefer to take a more passive role. For these individuals, it should be made clear that use of the app is optional. However, further work should also explore how to repackaging the content of MySurgery and other such interventions into a format that may empower and educate these individuals without pushing them outside of their comfort zone or placing a perceived additional burden on them at a time when they may already feel anxious. When patients become better informed and prepared for their surgery (eg, by being exposed to the information in the app), safety-related behaviors such as communication with health care workers should improve naturally [15].

These findings highlight an important point for app developers in the health care sector. Although there are clearly many benefits of apps in mobilizing knowledge, it is naïve to think that apps alone democratize access to information. As demonstrated here, there are cohorts of the population who, for a range of reasons, are unable or unwilling to use health apps, even when they have access to the technology. These barriers to use may only be revealed when examining closely the perceptions of users from a diverse range of backgrounds. It may not be possible to overcome every barrier to app use, and it may be that certain apps remain unsuitable for certain individuals. However, in this study, given the broad support for the MySurgery app, it was agreed by the participants that recognizing the presence of barriers to use should not prevent its development or future implementation but rather be used to enhance the implementation process by building strategies to improve accessibility. One approach that is likely to be important for encouraging use across all patient groups is the provision of support and encouragement for patient involvement from health care professionals themselves, particularly for those behaviors that patients perceive as potentially confrontational (eg, asking a health care professional if they have washed their hands before examining them). Indeed, the important role of health care professionals in promoting the use of MySurgery was a strong emergent theme in this study and is a well-established finding in the safety literature [32,34].

Once the suggested improvements have been made to MySurgery, the next step will be to trial implementation of the app on a small scale within 1 or 2 surgical departments. However, such interventions need to be initiated at the right time with the right tools to be effective; therefore, careful planning of the implementation strategy will be key. Several important points were raised in the discussions that should be fed into this approach. Regarding timing, it was agreed that users should be signposted toward the app just before their preoperative appointment when the information contained is most salient to them and when they have time to act on the advice, for example, by discussing safety-related matters with their clinician at their appointment and by ensuring they have followed the advice about preparing for surgery. It was also deemed critical to engage clinicians in the roll out of the app such that they are aware of the intervention and can promote it with their patients and empower them to discuss any concerns. This will likely involve not only education sessions and the appointment of clinical champions but also, as mentioned before, a long-term cultural effort to break down resistance to patient

involvement. Finally, part of the implementation approach should focus on thorough communication of the remit and scope of MySurgery, the broader safety context within which it sits, and available assistance for downloading and using the app.

### Strengths and Limitations

Common to qualitative research of this type, we had a small sample size, which limits the generalizability of the results. However, the objective of choosing focus groups over quantitative approaches is that it allows for the generation of far-richer data sets, which are desirable during the early phases of evaluating an intervention and planning its implementation. We achieved data saturation in the analysis, which allowed us to understand in depth the perceptions of the intervention and a theory of how it might work and the various themes we discussed relating to its accessibility and implementation. By carefully planning a diversity approach, we had representation from individuals from groups that tend to be underrepresented in health care research, including those with a disability, those from a BAME background, and those from the LGBT+ community. We feel this enriched the data in making it more representative of the population, encapsulating wide-ranging views of individuals from diverse backgrounds. This resulted in several points being raised, particularly around the inclusivity of the MySurgery intervention, which may not have been captured otherwise. To achieve this diversity, it was important to plan the focus group meetings carefully to ensure that they were accessible to all, that financial compensation for time and travel was provided, and that all dietary requirements were catered for. Incorporating a diversity approach in research such as this will allow us to have a more nuanced way of understanding the needs of different population groups and therefore in addressing if and where digital exclusions apply and how they can best be addressed. Given the close links between engagement with a digital intervention and adherence to its content, establishing an understanding of the factors that will influence engagement is a critical activity that should be undertaken with direct input from service users early on in the process of app design. This helps in identifying the design features that will draw users to (or deter them from) the intervention in the first place and features of the app that will enhance user motivation and autonomy and personal relevance and credibility of the intervention for different groups [20,35].

### Next Steps

Going forward with this program of research, it will be important to triangulate our findings with the implementation science literature to provide a theoretical lens and systematic approach to finalize the implementation plan. Implementation scientists are interested in understanding how best to promote the uptake of research findings into routine health care, calling on theoretical approaches to provide better understanding and explain how and why implementation succeeds or fails. Applying implementation science theory can help to map out the entire implementation approach—using a single taxonomy to identify the factors that might influence implementation effectiveness, including identifying the relevant stakeholders, the likely barriers to and facilitators of implementation, and how these interact within the context at hand. If used to evaluate

the initial stages of implementation, theoretical frameworks can help to produce findings to inform stakeholders on improvements to the intervention and its implementation. The literature also provides guidance in the selection of implementation strategies (ie, discrete methods or techniques used to enhance the adoption and sustainability of an intervention). Recent research has identified more than 70 discrete implementation strategies relevant to health care researchers, ranging from assessing readiness for change within an organization; providing supervision and training; tailoring the intervention to the specific context, right through planning marketing approaches; and considering incentive plans [36]. By consulting and cross-referencing these with the strategies identified as being important in this study and matching them to the barriers and facilitators identified, the approach to implementation becomes more robust, and strategies that may be helpful but may not have been considered can be built-in. In addition, by applying a theoretical framework to the evaluation of the implementation, it can aid in the implementation, which increases the efficacy of research and allows results to be more reliably generalized and built upon across future studies and contexts [37,38]. A later step will be to consider economic evaluation of the intervention to establish its role (if any) in cost saving. Such evaluation is rarely attempted, despite offering a more evidence-based assessment of the scalability, sustainability, and benefits of broader investment in such

technology tools [39,40]. Finally, as raised in the focus group discussion, it seems likely that there will be interest from providers in tailoring apps such as MySurgery to specific specialties or hospital units, to achieve better streamlining and integration of processes of care. Adaptation of interventions to specific contexts in this sense is an important principle of quality improvement work of this kind, not least to engender increased buy-in from staff. We will, therefore, be looking to collaborate with NHS Trusts to produce tailored versions of MySurgery for evaluation going forward. We will continue to build on the theory and practice of incorporating a diversity and inclusion approach into this study, as we believe, for reasons already mentioned, that adopting this approach could have significant benefits in making interventions more effective.

## Conclusions

This study was successful in establishing a diverse and inclusive stakeholder group to provide formative in-depth feedback on the MySurgery app. The app was received enthusiastically. It was endorsed as a powerful tool for enhancing patient empowerment in general and an appropriate approach to addressing well-established aims to involve patients and their relatives directly in maintaining patient safety. Various adaptations to the app should be made to make it more accessible to certain groups, which will involve the development of a comprehensive and multipronged implementation approach informed by diverse stakeholders.

---

## Acknowledgments

The authors would like to acknowledge the valued contributions made by all members of the focus groups and the time they committed to the project. SR, JO, and NS were supported by the National Institute for Health Research (NIHR) Applied Research Collaboration South London (ARC South London) at the King's College Hospital NHS Foundation Trust. SR (creator of MySurgery) is supported by the NIHR through a Knowledge Mobilization Fellowship to evaluate the MySurgery app over a 3-year program. JO was supported by the Health Foundation through an Improvement Science Fellowship. NS is a member of King's Improvement Science, which is part of the NIHR ARC South London and comprises a specialist team of improvement scientists and senior researchers based at the King's College London. Its work was funded by King's Health Partners (Guy's and St Thomas' NHS Foundation Trust, King's College Hospital NHS Foundation Trust, King's College London, and South London and Maudsley NHS Foundation Trust), Guy's and St Thomas' Charity, the Maudsley Charity, and the Health Foundation. 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.

---

## Conflicts of Interest

NS is the director of the London Safety and Training Solutions Ltd, which offers training in patient safety, implementation solutions, and human factors to health care organizations.

---

### Multimedia Appendix 1

Study advert.

[[DOCX File , 600 KB - mhealth\\_v9i4e24065\\_app1.docx](#) ]

---

### Multimedia Appendix 2

Equality monitoring form.

[[DOCX File , 114 KB - mhealth\\_v9i4e24065\\_app2.docx](#) ]

---

### Multimedia Appendix 3

Questionnaire regarding previous experience of surgery and smartphone apps.

[[DOCX File , 157 KB - mhealth\\_v9i4e24065\\_app3.docx](#) ]

## Multimedia Appendix 4

Focus group discussion guide.

[\[DOC File , 231 KB - mhealth\\_v9i4e24065\\_app4.doc \]](#)

## Multimedia Appendix 5

Illustrative quotes.

[\[DOCX File , 18 KB - mhealth\\_v9i4e24065\\_app5.docx \]](#)**References**

1. Ryu S. Book Review: mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). *Healthc Inform Res* 2012;18(3):231 [FREE Full text] [doi: [10.4258/hir.2012.18.3.231](https://doi.org/10.4258/hir.2012.18.3.231)]
2. Calvillo J, Román I, Roa LM. How technology is empowering patients? A literature review. *Health Expect* 2015 Oct;18(5):643-652 [FREE Full text] [doi: [10.1111/hex.12089](https://doi.org/10.1111/hex.12089)] [Medline: [23711169](https://pubmed.ncbi.nlm.nih.gov/23711169/)]
3. Alhomod A, Alzahrani S. Patient Empowerment via Mobile Personal Health Records and Mobile Health Applications: A Review of the Current Use. 2019 Aug Presented at: Portland International Conference on Management of Engineering and Technology (PICMET) (pp. ). IEEE; 2019 Aug 25; Portland, USA p. 1-4. [doi: [10.23919/picmet.2019.8893847](https://doi.org/10.23919/picmet.2019.8893847)]
4. All Party Parliamentary Groups. Patient empowerment: for better quality, more sustainable health services globally. London: House of Commons. 2014. URL: <https://www.iapo.org.uk/news/2014/may/22/report-patient-empowerment-better-quality-more-sustainable-health-services-globally> [accessed 2021-03-11]
5. The Joint Commission. Speak Up About Your Care. 2018. URL: <https://www.jointcommission.org/resources/for-consumers/speak-up-campaigns/about-your-care/> [accessed 2021-03-11]
6. Tritter J. Public and patient participation in health care and health policy in the United Kingdom. *Health Expect* 2011 Jun;14(2):220-223 [FREE Full text] [doi: [10.1111/j.1369-7625.2011.00697.x](https://doi.org/10.1111/j.1369-7625.2011.00697.x)] [Medline: [21592265](https://pubmed.ncbi.nlm.nih.gov/21592265/)]
7. 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](https://doi.org/10.1542/peds.2019-2056G)] [Medline: [32358210](https://pubmed.ncbi.nlm.nih.gov/32358210/)]
8. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. *Psychooncology* 2020 Jan;29(1):17-37. [doi: [10.1002/pon.5280](https://doi.org/10.1002/pon.5280)] [Medline: [31692183](https://pubmed.ncbi.nlm.nih.gov/31692183/)]
9. 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](https://doi.org/10.2196/mhealth.3335)] [Medline: [25803705](https://pubmed.ncbi.nlm.nih.gov/25803705/)]
10. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. *JAMA Intern Med* 2016 Mar;176(3):340-349. [doi: [10.1001/jamainternmed.2015.7667](https://doi.org/10.1001/jamainternmed.2015.7667)] [Medline: [26831740](https://pubmed.ncbi.nlm.nih.gov/26831740/)]
11. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. *J Med Internet Res* 2016 Oct 31;18(11):e287 [FREE Full text] [doi: [10.2196/jmir.5692](https://doi.org/10.2196/jmir.5692)] [Medline: [27806926](https://pubmed.ncbi.nlm.nih.gov/27806926/)]
12. Make + Ship. MySurgery iPhone app helps NHS patients get the best outcome from their surgery. URL: <https://www.makeandship.com/project/mysurgery/> [accessed 2021-03-11]
13. Harris K, Søfteland E, Moi AL, Harthug S, Storesund A, Jesuthasan S, et al. Patients' and healthcare workers' recommendations for a surgical patient safety checklist - a qualitative study. *BMC Health Serv Res* 2020 Jan 16;20(1):43 [FREE Full text] [doi: [10.1186/s12913-020-4888-1](https://doi.org/10.1186/s12913-020-4888-1)] [Medline: [31948462](https://pubmed.ncbi.nlm.nih.gov/31948462/)]
14. Davis R, Sevdalis N, Jacklin R, Vincent C. An examination of opportunities for the active patient in improving patient safety. *J Patient Saf* 2012 Mar;8(1):36-43. [doi: [10.1097/PTS.0b013e31823cba94](https://doi.org/10.1097/PTS.0b013e31823cba94)] [Medline: [22258225](https://pubmed.ncbi.nlm.nih.gov/22258225/)]
15. World Health Organisation. What you need to know before and after surgery. 2015. URL: [https://www.who.int/surgery/publications/patients\\_communication\\_tool.pdf?ua=1](https://www.who.int/surgery/publications/patients_communication_tool.pdf?ua=1) [accessed 2021-03-11]
16. Ljungqvist O, Scott M, Fearon KC. Enhanced Recovery After Surgery: A Review. *JAMA Surg* 2017 Mar 01;152(3):292-298. [doi: [10.1001/jamasurg.2016.4952](https://doi.org/10.1001/jamasurg.2016.4952)] [Medline: [28097305](https://pubmed.ncbi.nlm.nih.gov/28097305/)]
17. Howard R, Yin YS, McCandless L, Wang S, Englesbe M, Machado-Aranda D. Taking Control of Your Surgery: Impact of a Prehabilitation Program on Major Abdominal Surgery. *J Am Coll Surg* 2019 Jan;228(1):72-80 [FREE Full text] [doi: [10.1016/j.jamcollsurg.2018.09.018](https://doi.org/10.1016/j.jamcollsurg.2018.09.018)] [Medline: [30359831](https://pubmed.ncbi.nlm.nih.gov/30359831/)]
18. Khan SK, Malviya A, Muller SD, Carluke I, Partington PF, Emmerson KP, et al. Reduced short-term complications and mortality following Enhanced Recovery primary hip and knee arthroplasty: results from 6,000 consecutive procedures. *Acta Orthop* 2014 Feb;85(1):26-31 [FREE Full text] [doi: [10.3109/17453674.2013.874925](https://doi.org/10.3109/17453674.2013.874925)] [Medline: [24359028](https://pubmed.ncbi.nlm.nih.gov/24359028/)]
19. Russ S, Latif Z, Hazell AL, Ogunmuyiwa H, Tapper J, Wachuku-King S, et al. A Smartphone App Designed to Empower Patients to Contribute Toward Safer Surgical Care: Community-Based Evaluation Using a Participatory Approach. *JMIR Mhealth Uhealth* 2020 Jan 20;8(1):e12859 [FREE Full text] [doi: [10.2196/12859](https://doi.org/10.2196/12859)] [Medline: [31958067](https://pubmed.ncbi.nlm.nih.gov/31958067/)]

20. 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* 2017 Jun;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
21. Lai P. THE LITERATURE REVIEW OF TECHNOLOGY ADOPTION MODELS AND THEORIES FOR THE NOVELTY TECHNOLOGY. *JISTEM* 2017 Apr 30;14(1). [doi: [10.4301/s1807-17752017000100002](https://doi.org/10.4301/s1807-17752017000100002)]
22. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' Acceptance of Smartphone Health Technology for Chronic Disease Management: A Theoretical Model and Empirical Test. *JMIR Mhealth Uhealth* 2017 Dec 06;5(12):e177 [FREE Full text] [doi: [10.2196/mhealth.7886](https://doi.org/10.2196/mhealth.7886)] [Medline: [29212629](https://pubmed.ncbi.nlm.nih.gov/29212629/)]
23. Hennings C, Herstatt C. Belief elicitation study: Identifying salient beliefs of patients towards the use of mHealth. Working Paper; 2019. URL: <https://www.econstor.eu/handle/10419/192994> [accessed 2012-03-11]
24. Tritter J, Landstad B. Focus Groups. In: *Qualitative Research in Health Care*. Online: John Wiley & Sons Ltd; Jan 2020:15-66.
25. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. *BMJ Qual Saf* 2016 Aug;25(8):626-632 [FREE Full text] [doi: [10.1136/bmjqs-2015-004839](https://doi.org/10.1136/bmjqs-2015-004839)] [Medline: [26993640](https://pubmed.ncbi.nlm.nih.gov/26993640/)]
26. Ocloo J. Being heard, not, "seldom heard": democratising research with diverse communities during the covid-19 pandemic. *BMJ Opinion*. 2020. URL: <https://blogs.bmj.com/bmj/2020/06/02/being-heard-not-seldom-heard-democratising-research-with-diverse-communities-during-the-covid-19-pandemic/> [accessed 2021-03-11]
27. Equality Act 2010. Equality Act. 2010. URL: <https://www.legislation.gov.uk/ukpga/2010/15/contents> [accessed 2021-03-11]
28. Dawson S, Campbell SM, Giles SJ, Morris RL, Cheraghi-Sohi S. Black and minority ethnic group involvement in health and social care research: A systematic review. *Health Expect* 2018 Feb;21(1):3-22 [FREE Full text] [doi: [10.1111/hex.12597](https://doi.org/10.1111/hex.12597)] [Medline: [28812330](https://pubmed.ncbi.nlm.nih.gov/28812330/)]
29. NIHR. Going the extra mile: Improving the nation's health and well-being through public involvement in research. URL: <https://www.nihr.ac.uk/documents/about-us/our-contribution-to-research/how-we-involve-patients-carers-and-the-public/Going-the-Extra-Mile.pdf> [accessed 2021-03-11]
30. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation* 2016 Jun 30;27(2):237-246. [doi: [10.1177/1098214005283748](https://doi.org/10.1177/1098214005283748)]
31. Williams R. One in five people aged 75 and over now use a smartphone. Jan. 2019. URL: <https://inews.co.uk/news/one-in-five-people-aged-75-and-over-now-use-a-smartphone-245856> [accessed 2021-03-11]
32. Davis RE, Sevdalis N, Vincent CA. Patient involvement in patient safety: How willing are patients to participate? *BMJ Qual Saf* 2011 Jan;20(1):108-114. [doi: [10.1136/bmjqs.2010.041871](https://doi.org/10.1136/bmjqs.2010.041871)] [Medline: [21228083](https://pubmed.ncbi.nlm.nih.gov/21228083/)]
33. The Guardian. Digital divide 'isolates and endangers' millions of UK's poorest. 2020. URL: <https://www.theguardian.com/world/2020/apr/28/digital-divide-isolates-and-endangers-millions-of-uk-poorest> [accessed 2021-03-11]
34. Rose SA, Poynter PS, Anderson JW, Noar SM, Conigliaro J. Physician weight loss advice and patient weight loss behavior change: a literature review and meta-analysis of survey data. *Int J Obes (Lond)* 2013 Jan;37(1):118-128. [doi: [10.1038/ijo.2012.24](https://doi.org/10.1038/ijo.2012.24)] [Medline: [22450855](https://pubmed.ncbi.nlm.nih.gov/22450855/)]
35. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Med Inform Decis Mak* 2017 Feb 28;17(1):25 [FREE Full text] [doi: [10.1186/s12911-017-0422-8](https://doi.org/10.1186/s12911-017-0422-8)] [Medline: [28241759](https://pubmed.ncbi.nlm.nih.gov/28241759/)]
36. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci* 2015 Feb 12;10:21 [FREE Full text] [doi: [10.1186/s13012-015-0209-1](https://doi.org/10.1186/s13012-015-0209-1)] [Medline: [25889199](https://pubmed.ncbi.nlm.nih.gov/25889199/)]
37. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. *Am J Prev Med* 2012 Sep;43(3):337-350 [FREE Full text] [doi: [10.1016/j.amepre.2012.05.024](https://doi.org/10.1016/j.amepre.2012.05.024)] [Medline: [22898128](https://pubmed.ncbi.nlm.nih.gov/22898128/)]
38. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci* 2015 Apr 21;10:53 [FREE Full text] [doi: [10.1186/s13012-015-0242-0](https://doi.org/10.1186/s13012-015-0242-0)] [Medline: [25895742](https://pubmed.ncbi.nlm.nih.gov/25895742/)]
39. 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](https://doi.org/10.2196/mhealth.6425)] [Medline: [27780795](https://pubmed.ncbi.nlm.nih.gov/27780795/)]
40. 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. [doi: [10.1371/journal.pone.0170581](https://doi.org/10.1371/journal.pone.0170581)] [Medline: [28152012](https://pubmed.ncbi.nlm.nih.gov/28152012/)]

## Abbreviations

- BAME:** Black, Asian, and minority ethnic  
**CIS:** Centre for Implementation Science  
**LGBT+:** lesbian, gay, bisexual, transgender, and queer

**NHS:** National Health Service  
**NIHR:** National Institute for Health Research  
**PPI:** Patient and Public Involvement

*Edited by G Eysenbach; submitted 15.09.20; peer-reviewed by S Badawy, K Ferrini; comments to author 31.10.20; revised version received 15.01.21; accepted 31.01.21; published 08.04.21.*

*Please cite as:*

Russ S, Sevdalis N, Ocloo J

*A Smartphone App Designed to Empower Patients to Contribute Toward Safer Surgical Care: Qualitative Evaluation of Diverse Public and Patient Perceptions Using Focus Groups*

*JMIR Mhealth Uhealth* 2021;9(4):e24065

URL: <https://mhealth.jmir.org/2021/4/e24065>

doi: [10.2196/24065](https://doi.org/10.2196/24065)

PMID: [33830062](https://pubmed.ncbi.nlm.nih.gov/33830062/)

©Stephanie Russ, Nick Sevdalis, Josephine Ocloo. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 08.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Measuring the Quality of Clinical Skills Mobile Apps for Student Learning: Systematic Search, Analysis, and Comparison of Two Measurement Scales

Tehmina Gladman<sup>1</sup>, PhD; Grace Tylee<sup>2</sup>, MBChB; Steve Gallagher<sup>3</sup>, PhD; Jonathan Mair<sup>2</sup>, MBChB; Rebecca Grainger<sup>1</sup>, PhD

<sup>1</sup>Education Unit, University of Otago Wellington, Wellington, New Zealand

<sup>2</sup>Hutt Valley District Health Board, Lower Hutt, New Zealand

<sup>3</sup>Education Unit, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

**Corresponding Author:**

Tehmina Gladman, PhD

Education Unit

University of Otago Wellington

PO Box 7343

Newtown

Wellington, 6242

New Zealand

Phone: 64 04 918 6749

Email: [tehmina.gladman@otago.ac.nz](mailto:tehmina.gladman@otago.ac.nz)

## Abstract

**Background:** Mobile apps are widely used in health professions, which increases the need for simple methods to determine the quality of apps. In particular, teachers need the ability to curate high-quality mobile apps for student learning.

**Objective:** This study aims to systematically search for and evaluate the quality of clinical skills mobile apps as learning tools. The quality of apps meeting the specified criteria was evaluated using two measures—the widely used Mobile App Rating Scale (MARS), which measures general app quality, and the Mobile App Rubric for Learning (MARuL), a recently developed instrument that measures the value of apps for student learning—to assess whether MARuL is more effective than MARS in identifying high-quality apps for learning.

**Methods:** Two mobile app stores were systematically searched using clinical skills terms commonly found in medical education and apps meeting the criteria identified using an approach based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. A total of 9 apps were identified during the screening process. The apps were rated independently by 2 reviewers using MARS and MARuL.

**Results:** The intraclass correlation coefficients (ICCs) for the 2 raters using MARS and MARuL were the same (MARS ICC [two-way]=0.68;  $P<.001$  and MARuL ICC [two-way]=0.68;  $P<.001$ ). Of the 9 apps, Geeky Medics-OSCE revision (MARS Android=3.74; MARS iOS=3.68; MARuL Android=75; and MARuL iOS=73) and OSCE PASS: Medical Revision (MARS Android=3.79; MARS iOS=3.71; MARuL Android=69; and MARuL iOS=73) scored highly on both measures of app quality and for both Android and iOS. Both measures also showed agreement for the lowest rated app, Patient Education Institute (MARS Android=2.21; MARS iOS=2.11; MARuL Android=18; and MARuL iOS=21.5), which had the lowest scores in all categories except information (MARS) and professional (MARuL) in both operating systems. MARS and MARuL were both able to differentiate between the highest and lowest quality apps; however, MARuL was better able to differentiate apps based on teaching and learning quality.

**Conclusions:** This systematic search and rating of clinical skills apps for learning found that the quality of apps was highly variable. However, 2 apps—Geeky Medics-OSCE revision and OSCE PASS: Medical Revision—rated highly for both versions and with both quality measures. MARS and MARuL showed similar abilities to differentiate the quality of the 9 apps. However, MARuL's incorporation of teaching and learning elements as part of a multidimensional measure of quality may make it more appropriate for use with apps focused on teaching and learning, whereas MARS's more general rating of quality may be more appropriate for health apps targeting a general health audience. Ratings of the 9 apps by both measures also highlighted the variable quality of clinical skills mobile apps for learning.

**KEYWORDS**

mobile apps; MARS; MARuL; medical education; app review; mobile phone

## Introduction

### Background

Mobile apps are widely used by health care professionals and have been shown to improve documentation, workflows, access to information, and clinical decision support [1]. Apps can be found from web-based vendors (app stores), web-based repositories (app repositories, online communities, and news stories), and peer-reviewed literature [2]. A recently published framework for finding apps [3] recommends peer-reviewed literature as the first source of information on quality apps. This nascent body of literature includes high-quality evaluations of single apps and systematic searches of app stores for apps, often including the appraisal of app quality. There is an emerging literature on systematic app store searches for apps to support clinical care [4-10] and the development of instruments for assessing app quality, such as the Mobile App Rating Scale (MARS) [11]. Although mobile apps are now widely accessible and being implemented in clinical care, the role of mobile apps in medical education is less well evaluated.

Mobile app use for teaching and learning has been an area of exploration since the first smartphones became available, and there has been growing use since then [12]. With this increased use, some studies have aimed to determine the characteristics of mobile apps that best contribute to student learning [12,13]. The app characteristics that users identify as best-promoting self-regulated and deep learning include perceived usefulness, perceived satisfaction, and interactivity [13,14]. Although frameworks for implementing mobile technology in medical education have been proposed [15], the evaluation of mobile technology use among medical students is largely limited to surveys evaluating types of apps used and extent of use [16-19] and barriers and facilitators to the use of mobile devices [20,21]. To date, there have not been many studies to identify or evaluate apps to support medical student learning or any systematic app store searches to identify and evaluate the potential quality of apps aimed at medical students. Such studies would be useful for medical teachers in their role as resource curators [22] so that they can easily compare, identify, and direct students to content-relevant, high-quality apps to support learning. Medical students would also be consumers of such research to find apps that may support self-directed learning. By considering aspects of an app, such as the usefulness of the content being presented, the interactivity of the app in its presentation of content, and its use of methods of learning that increase student satisfaction and interest, such as case-based learning [23], and combining user-centered qualities with technology-centered qualities such as functionality, stability, esthetic appeal, and ease of use [24], we can identify apps that are likely to be effective aids for learning.

We have previously worked with medical students to develop a rubric to evaluate the value of mobile apps to support medical student just-in-time learning [25]. This instrument, the Mobile

App Rubric for Learning (MARuL), can be rapidly and easily used by teachers or students to rate the quality of an app and its potential to be useful for learning. MARuL contains 4 categories: teaching and learning measures (n=9), user-centered measures (n=7), professional measures (n=3), and usability measures (n=7). As mobile apps do not yet seem to be widely endorsed or promoted by medical schools to support learning [15,26], MARuL may offer a tool for the faculty to confidently evaluate the quality of apps to support learning [27].

Although the general quality of any health app can be evaluated with the well-established MARS instrument, apps for medical student learning are a subset of health apps that have a specific purpose requiring additional aspects for evaluation. MARuL, though adapting 9 items from MARS, was designed specifically to measure aspects of an app related to its value for medical student learning [25].

### Objectives

This study reports on the use of MARS and MARuL to evaluate apps designed to help medical students develop clinical skills. Clinical skills are a competency that all medical students need to acquire, requiring complex knowledge, psychomotor skills, and integration skills. Good-quality apps could be a useful learning tool for students to acquire these skills. We define clinical skills as any discrete and observable act within the overall process of patient care [27], and for the purposes of this study, we focus on clinical skills required during a traditional doctor-patient interaction. The apps of interest might support the development of history taking, physical examination skills, and patient explanation, which are often assessed in objective structured clinical examinations (OSCEs).

To extend previous work in developing methods of systematic app store search and app evaluation [28] specifically for apps for medical student learning, we aim to do the following:

- Undertake a systematic search of app stores to identify apps available to support clinical skills development by medical students.
- Evaluate the perceived quality of those apps using MARS and the potential value of those apps for student just-in-time learning of clinical skills using MARuL.
- Compare MARS and MARuL as methods for evaluating perceived quality and value of apps for learning.

## Methods

### App Identification

We performed a systematic app search in the New Zealand Apple iOS App Store and Google Play Store between January 15 and February 1, 2019. Search terms were chosen to focus on apps for teaching and learning in health. Three of the authors (TG, GT, and RG) developed the search terms and inclusion and exclusion criteria through preliminary searches and



discussion. The final set contained 14 search terms that were searched one at a time ([Textbox 1](#)).

**Textbox 1.** Search terms used in app stores grouped by focus of search term.

<p>General:</p> <ul style="list-style-type: none"> <li>• Clinical skills</li> </ul> <p>Objective Structured Clinical Examination:</p> <ul style="list-style-type: none"> <li>• OSCE</li> <li>• Objective Structured Clinical Examination</li> </ul> <p>History taking:</p> <ul style="list-style-type: none"> <li>• Medical history taking</li> <li>• Clinical history taking</li> <li>• Patient history</li> </ul> <p>Examinations:</p> <ul style="list-style-type: none"> <li>• Medical examination</li> <li>• Medical exam</li> <li>• Physical examination</li> <li>• Physical exam</li> <li>• Clinical examination</li> <li>• Clinical exam</li> </ul> <p>Explanation:</p> <ul style="list-style-type: none"> <li>• Planning and explaining</li> <li>• Patient education</li> </ul>
---

## Eligibility Criteria

Apps were initially screened during the search by reading the title and description of the app in the app store. Apps were eligible for inclusion in the review if, in the initial screening, they fulfilled the following 6 inclusion criteria: (1) were available in English; (2) included at least one of the keywords ([Textbox 1](#)) in the title or description; (3) included an interactive element requiring some form of input (as deliberate practice with active learning is more effective [14]—to be interactive, an app must require students to perform in some way, eg, by filling in a form, answering questions, or interacting with an image by rotation or other means); (4) their target audience included medical students based on a statement in the app description; (5) supported iOS 8 or later and Android version 5 or later (to include devices in the last 5 years that used these systems); (6) were available for both Android and iOS to ensure student accessibility.

Apps were excluded if they failed to meet the inclusion criteria or if they met any of the following exclusion criteria: (1) priced more than NZ \$10 (US \$7) for a monthly subscription or as a one-off price; based on a discussion with GT, student research collaborator, and RG, a local leader in medical education and experienced clinician; this was thought to be a reasonable maximum cost that either a student would spend on themselves or an institution would be willing to spend per user; (2) were reference-only apps (passive with no student input, ie, do not

require students to interact beyond basic touchscreen requirements such as page turning or pressing play on a video, eg, textbook apps or apps that contain videos of clinical skills being performed that students watch but do not interact with); (3) designed for staff-only use in formative or summative assessment contexts; (4) complemented other software (not stand alone); and (5) required a log in or sign up to be used [29,30] based on a discussion with GT who noted that requiring an initial signup or registration was a barrier to use for most of her student colleagues. These exclusion criteria were based on potential barriers to students' use or reduced quality of learning for students.

## Data Extraction

A data screening and extraction spreadsheet was developed and refined by 2 researchers (GT and TG) using [Airtable](#) [31] before the search. The app name, developer, operating system, reviewer, and whether the app was included or excluded were recorded in the spreadsheet during the initial search and screening of the app store search. Apps were excluded if one of the exclusion criteria was met, and the reason for exclusion was recorded. The iOS store was searched using an iPhone 7 (Apple Inc) and an iPhone 8 using iOS version 12.1.3, and the Google Play Store was searched using 2 Samsung Galaxy J1 Ace phones using Android version 5.1.1.

The app search for the iOS App Store and Google Play Store was completed in parallel but independently by 2 authors (GT

and TG). GT and TG then jointly reviewed the apps where there was a lack of agreement. Each discrepancy was discussed, and a final decision was made regarding inclusion or exclusion and the grounds for exclusion. The search and screening are reported based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) phase 1 guidelines, modified for app stores and app metadata [32].

### App Rating

All included apps were independently rated by 2 reviewers (JM and SG) with MARS [11] and MARuL [25]. The 2 reviewers were chosen because of their relationship to the student experience. One reviewer is a near peer of medical students, and the other reviewer works extensively with web-based learning for medical education to support student learning. Both MARS and MARuL versions include instructions to consider the target audience for the app, and the individual items of each measure use language keeping the target audience in mind.

First, the 2 reviewers met on videoconference to confirm their understanding of each rubric and its submeasures. They then completed a pilot rating on one excluded app and met on videoconference to discuss their scoring on the items and come to an agreement on how to interpret items that they differed on. The reviewers then independently downloaded and reviewed the included apps in iOS (iPhone 6 Plus and iPhone 6s) and Android (Samsung Galaxy J1 Ace) between November 10 and December 9, 2019. App reviews were completed using the MARS and MARuL in a web-based form (Qualtrics), with data exported to an Excel spreadsheet (Microsoft Office, version 16.40.20081000). The reviewers interacted with each app to fully explore its features before completing the MARS and MARuL. Both category and overall scores on MARS and MARuL for each app were calculated for each reviewer. To measure interrater reliability for MARS and MARuL, intraclass correlation coefficient (ICC) estimates were calculated with their 95% CIs in RStudio [33] based on a single-rating, consistency, two-way mixed effects model [34].

MARS comprises four categories of perceived app quality—engagement, functionality, esthetics, and information—and 1 category of subjective quality. Each category score is the mean of the items, rated on a 5-point Likert-type scale (from 1=inadequate to 5=excellent) within its category. The overall quality score was calculated by taking the mean of the 4 app quality category scores, with a final score ranging from 0 to 5 [11].

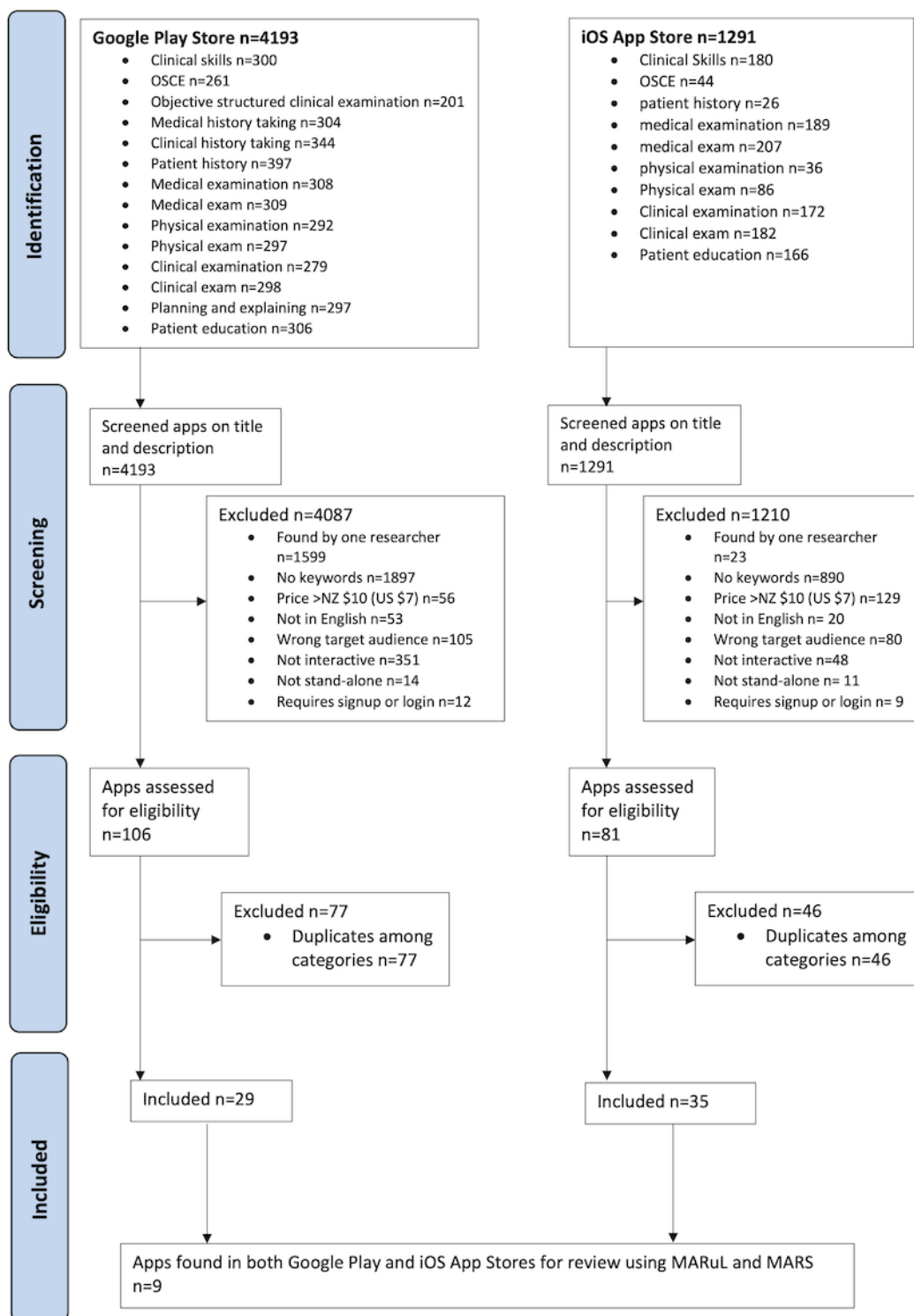
MARuL is composed of four categories, each of which receives a score. The category scores are summed together to reach an overall value for the learning score. The MARuL category score is calculated by adding the rating for each item on a 5-point Likert-type scale (0=does not fulfill the item requirements, 1=poorly fulfills the requirements, 2=somewhat fulfills the requirements, 3=mostly fulfills the requirements, and 4=fully meets the requirements) within each category to reach a total score for that category (teaching and learning=36, user-centered=28, professional measures=12, and usability=28). Summing the categories gives the user an overall score of 104. Apps are then categorized by their scoring range (<50=not at all valuable, 51-69=potentially valuable, and >69=probably valuable) [25].

## Results

### App Store Search

A total of 1291 iOS apps and 4193 Android apps were screened in the iOS App Store and Google Play Store, respectively. Following the app title and description screening, 1210 iOS apps and 4087 Android apps were excluded. Despite using the same search terms, sometimes discrepancies in the results from the search carried out by the 2 reviewers were seen, such as some apps only being found by one reviewer. Only the apps that were found by both reviewers were included in the final sample. We made this decision as our goal for this study was to rate commonly found apps available in both iOS and Android app stores. If the same apps appeared in a search by both a student and staff member, it was felt that they would be commonly located despite any search optimization in use. The two main reasons for exclusion of Android apps were that no keywords were found within the title or description (1897/4087, 46.4%) or that they were only found by one of the 2 researchers (1599/4087, 39.1%). The two main reasons for exclusion of iOS apps were no keywords in the title or description (890/1210, 73.6%), followed by a price greater than NZD \$10 (US \$7) as a one-off or recurring cost (129/1210, 10.7%). For iOS, 81 apps from the 14 search terms were identified, 35 of which were unique apps. For Android apps, a total of 106 apps were identified, of which 29 were unique. Of the 35 unique iOS apps and 29 unique Android apps, 9 apps were found on both iOS App Store and Google Play Store for inclusion. A search of the Apple Store in the United States using the website [fnd.io](http://fnd.io) [35] did not find any further apps that were also available in the international Google Play Store. [Figure 1](#) shows the search and screening process.

**Figure 1.** Flowchart for the identification of the Google Play Store and iOS App Store clinical skills apps. MARS: Mobile App Rating Scale; MARuL: Mobile App Rubric for Learning. OSCE: objective structured clinical examinations.



### App Characteristics

The characteristics of the apps are summarized in Table 1. The apps ranged in size from 2.7 to 229.2 MB. All apps were free to download except OSCE PASS: Medical Revision, which cost

NZD \$10 (US \$7). Three of the apps, Geeky Medics-OSCE revision; InSimu: The Patient Simulator; and Resuscitation!, had in-app purchases available for additional content. Of the 9 apps, 7 apps were stated to be for medical students and 5 apps specifically focused on clinical skills for OSCE practice.

**Table 1.** Characteristics of the 9 included apps.

App name	Developer	App version	Cost	App size (MB)	Description
CardioSmart; Heart Explorer	American College of Cardiology	3.0 (iOS); 2.3 (Android)	Free	229.2	Interactive 3D heart anatomy and educational pathology videos
Geeky Medics-OSCE <sup>a</sup> revision	Geeky Medics LTD	2.81 (iOS); 2.46 (Android)	Free; in-app purchases up to NZ \$18.99 (US \$13.41) on iOS and NZ \$29.99 (US \$21.18) on Android	76.6	OSCE guides for medical students
InSimu—The Patient Simulator (iOS); InSimu—Diagnose Virtual Clinical Cases (Android)	InSimu	1.7.7 (iOS); 1.8.7 (Android)	Free; in-app purchases up to NZ \$499.99 (US \$353.09) for lifetime access	51.8	Virtual clinic environment or simulation, work-through diagnosis
OSCE PASS: Medical Revision	Entremed Ltd	1.1 (iOS); 1.0 (Android)	NZ \$10 (US \$7)	8.3 (iOS); 3.79 (Android)	Written guides and video demonstrations for clinical skills
OSCE Revision for Medical Students (iOS); OSCE Revision (Android)	Matthew Roche	1.2.1 (iOS); 1.1.5 (Android)	Free	10.2	OSCE revision guides with test function
OSCEr	Ahmad Alhashemi (iOS); Essentials of clinical examination (Android)	1.0	Free	22.2	Guides for clinical skills with test and practice options
Pocket PEX: Physical Exam Aid (iOS); Pocket PEX (Android)	Charles Goldberg (iOS); MedEd Apps (Android)	3.1	Free	2.9	Interactive checklists for physical examination
Resuscitation!	EM Gladiators LLC	2.8 (iOS); 2.0 (Android)	Free; in-app purchases up to NZ \$16.99 (US \$11.99)	82.7	Virtual patient simulator, work-through diagnosis
Patient Education Institute	Olaf Breukhoven (iOS); The Patient Education Institute (Android)	1.2.3 (iOS); 1.2.2 (Android)	Free	2.7	Medical illustrations

<sup>a</sup>OSCE: objective structured clinical examination.

## App Rating

The 9 apps reviewed by the 2 researchers were CardioSmart Heart Explorer; Geeky Medics-OSCE revision; InSimu—The Patient Simulator; OSCE Revision; OSCEr; Pocket PEX;

Physical Exam Aid; OSCE PASS: Medical Revision; Patient Education Institute; and Resuscitation! ICC scores for MARuL was ICC (two-way)=0.68 ( $P<.001$ ) and for MARS was ICC (two-way)=0.68 ( $P<.001$ ), indicating moderate reliability (Table 2) [34].

**Table 2.** Interrater reliability scores for the Mobile App Rubric for Learning and Mobile App Rating Scale.

Rating measures	Intraclass correlation	95% CI	<i>F</i> test with true value 0	
			<i>F</i> test ( <i>df</i> )	<i>P</i> value
MARS <sup>a</sup>	0.677	0.618-0.729	5.2 (367)	<.001
MARuL <sup>b</sup>	0.676	0.621-0.725	5.18 (415)	<.001

<sup>a</sup>MARS: Mobile App Rating Scale.

<sup>b</sup>MARuL: Mobile App Rubric for Learning.

The total app quality mean scores from the MARS evaluation ranged from 2.11 to 3.71 on the 9 iOS apps and 2.21 to 3.79 on the 9 Android apps (Table 3), with lowest scores generally occurring in the engagement and information categories. OSCE

PASS: Medical Revision (iOS=3.71; Android=3.79), Geeky Medics-OSCE revision (iOS=3.68; Android=3.74), and CardioSmart Heart Explorer (iOS=3.53; Android=3.53) were the top-scoring apps on iOS and Android.

**Table 3.** Average Mobile App Rating Scale scores from the 2 raters for the 9 apps tested.

Operating system and app name	Total	Engagement	Functionality	Aesthetics	Information	Subjective quality
<b>Android</b>						
OSCE <sup>a</sup> PASS: Medical Revision	3.79	3.60	4.63	3.83	3.43	3.88
Geeky Medics-OSCE revision	3.74	3.70	4.38	4.00	3.29	3.88
CardioSmart Heart Explorer	3.53	3.10	4.38	4.33	3.00	2.00
Resuscitation!	3.45	3.60	4.38	3.50	2.79	3.63
OSCEr	3.24	2.90	4.25	3.67	2.71	2.00
InSimu—The Patient Simulator	3.03	3.20	3.63	3.17	2.50	1.88
Pocket PEx: Physical Exam Aid	2.76	2.10	4.00	2.67	2.57	1.88
OSCE Revision for Medical Students	2.39	2.00	2.75	3.00	2.21	1.63
Patient Education Institute	2.21	1.60	2.75	2.17	2.36	1.00
<b>iOS</b>						
OSCE PASS: Medical Revision	3.71	3.40	4.50	4.00	3.36	3.88
Geeky Medics-OSCE revision	3.68	3.10	4.75	4.17	3.29	3.75
CardioSmart Heart Explorer	3.53	3.00	4.25	4.33	3.14	2.38
Resuscitation!	3.50	3.60	4.50	3.67	2.79	3.63
OSCEr	3.16	2.90	4.25	3.50	2.57	2.13
Pocket PEx: Physical Exam Aid	2.84	2.10	4.13	2.67	2.71	1.75
InSimu—The Patient Simulator	2.76	2.80	3.38	3.33	2.14	1.75
OSCE Revision for Medical Students	2.66	2.70	3.13	3.00	2.21	1.88
Patient Education Institute	2.11	1.50	2.63	2.00	2.29	1.00

<sup>a</sup>OSCE: objective structured clinical examination.

The MARuL overall app scores ranged from 21.5 to 73.0 for the 9 iOS apps and 18.0 to 75.0 for the 9 Android apps. Two apps, Geeky Medics-OSCE revision and OSCE PASS: Medical Revision, scored as *probably valuable* in both iOS and Android, and 1 app—Resuscitation!—as *potentially valuable* in both iOS

and Android (Table 4). CardioSmart Heart Explorer scored at the low end of the range for *potentially valuable* in Android only. The remaining apps had a MARuL score of less than 50 or *not at all valuable*.

**Table 4.** Average Mobile App Rubric for Learning scores from the 2 raters for the 9 apps tested.

Operating system and app name	Total score out of 104	User-centered score out of 28	Teaching and learning score out of 36	Professional score out of 12	Usability score out of 28
<b>Android</b>					
Geeky Medics-OSCE <sup>a</sup> revision	75	20	24	8.5	22.5
OSCE PASS: Medical Revision	69	20	24	5.5	19.5
Resuscitation!	65	19	21	5	20
CardioSmart Heart Explorer	54.5	11	12.5	9.5	21.5
OSCE Revision for Medical Students	50	8.5	19.5	6.5	15.5
OSCEr	47.5	9	15	6	17.5
Pocket PEx: Physical Exam Aid	46	7.5	14.5	8.5	15.5
InSimu—The Patient Simulator	40.5	8.5	8.5	5.5	18
Patient Education Institute	18	1	2	5	10
<b>iOS</b>					
Geeky Medics-OSCE revision	73	19.5	24	8.5	21
OSCE PASS: Medical Revision	73	19.5	25.5	8	20
Resuscitation!	66	18.5	22	5.5	20
CardioSmart Heart Explorer	49.5	10.5	10.5	10	18.5
Pocket PEx: Physical Exam Aid	49	8.5	16	8	16.5
OSCEr	48	9	16	6	17
InSimu—The Patient Simulator	45	9	8.5	7.5	20
OSCE Revision for Medical Students	39.5	5.5	14	6.5	13.5
Patient Education Institute	21.5	1.5	2	6	12

<sup>a</sup>OSCE: objective structured clinical examination.

## Discussion

### Principal Findings

This systematic app store search of the iOS App Store and Google Play Store for apps supporting the development of clinical skills required in the doctor-patient consultation in medical students resulted in the inclusion of 9 relevant apps. The evaluation of the 9 apps—using MARS [11] and MARuL [25]—found only 2 apps that scored highly in fulfilling the quality criteria across both measures of perceived quality for both mobile operating systems, Geeky Medics-OSCE revision and OSCE PASS: Medical Revision. However, each operating system and quality measure identified 3 apps that scored highly in fulfilling the criteria. The top 3 apps as rated by MARS were OSCE PASS: Medical Revision, Geeky Medics-OSCE revision, and CardioSmart Heart Explorer. For MARuL, Geeky Medics-OSCE revision, OSCE PASS: Medical Revision, and Resuscitation! were the top-scoring apps.

MARS and MARuL were designed to measure the perceived quality of apps for different purposes. Although MARS was developed as a method for measuring the perceived quality of a health mobile app for general use purposes [11], MARuL was specifically developed as a measure of the perceived value of a health education app to support student learning [25]. Both measures differentiated between apps of varying quality, as

shown by the similarity of their top-ranked apps and the consistency with which they categorized the lowest ranked app, Patient Education Institute, across most of their categories. The similar results for the ranking of apps across the 2 measures indicate that both measures are helpful in characterizing the perceived value or quality of mobile health apps. However, having a category specifically designed to measure teaching and learning allows teachers to use MARuL to measure perceived value for student learning. For example, although CardioSmart Heart Explorer was the third highest rated app using MARS, MARuL rated it at the low end of the category *potentially valuable* on Android devices (54.5), and the iOS version had an overall score of <50 (49.5). Scrutiny of the individual categories of MARuL reveals that CardioSmart Heart Explorer had the third lowest score for teaching and learning in both iOS (10.5) and Android (12.5).

After review using both MARS and MARuL, it was found that the quality of the 9 apps was highly variable. For example, in the MARS evaluations, apps tended to score the highest in the functionality category, followed by esthetics. The scores for engagement and information were the lowest. The engagement category of MARS considers whether the app is fun, interesting, customizable, interactive, and well-targeted to the audience. Similarly, the user-centered category of MARuL considers aspects of the app, such as satisfaction, user experience, and engagement. Of the 9 apps, 6 scored less than half of the

possible points in this category. It is concerning that apps consistently scored low in this category, as interest and enjoyment have been found to be strong influencers on students' persistence of learning [36-38].

One of the criteria for inclusion of apps in our review was the presence of interactivity within the app. Interactivity was evaluated in both MARS and MARuL. Interactivity increases engagement and may stimulate better learning of a topic [14]. Although all the apps reviewed were interactive, the degree of interactivity varied among apps. For example, the app Pocket PEx: Physical Exam Aid had minimal interactivity with checkboxes for each component of a physical exam, whereas the app InSimu had a comprehensive diagnostic scenario with interactivity for each step of the diagnostic process.

Both MARS and MARuL contain a category that includes items on information quality and credibility. Apps performed poorly in this category with only one app, Pocket PEx: Physical Exam Aid, containing easy-to-find references for information. No references were provided for the other apps. This poses a challenge for all types of health apps because of the importance of accurate and evidence-based information [39-41].

As noted earlier, MARuL has a category for the teaching and learning aspects of an app. It includes items on app purpose, pedagogy, capacity to generate learning, quantity of information, relevance to study or course, instructional features, user interactivity, feedback, and efficiency. The highest scoring apps in the teaching and learning category were Geeky Medics-OSCE revision, OSCE PASS, and Resuscitation!, which were also the top-scoring apps overall. Although the teaching and learning category has the highest weighting in the MARuL overall score, and the scoring trend for most apps across the other categories was similar to the teaching and learning category, taking a multidimensional approach to evaluation is important because of the interdependence of the dimensions in measuring value [24]. These findings of variable quality of clinical skills learning apps are consistent with findings from app reviews for patient-centered health-related apps [10,42] and are likely because of the poorly regulated market for mobile health apps.

### Limitations

The app store search was conducted in the New Zealand iOS App Store and Google Play Store. Although this could limit the

generalizability of our findings to other countries with different app stores, it should be noted that the Google Play Store is international and a search of the iOS store in the United States using [fnd.io](http://fnd.io) [35] did not find any new apps that were included in the Google Play Store. This limitation has been discussed in other app reviews. However, this study specifically focuses on New Zealand medical students; therefore, generalizability is not an immediate concern [43]. The app stores were searched in early 2019. As the rate of change in the app stores is high, it is possible that the apps we originally excluded have now changed enough to be included and other apps may have been removed since the search and review were conducted. The constantly changing nature of apps and their availability in app stores have also been discussed in previous reviews [44,45]. As such, it may prove challenging to keep an up-to-date list of good-quality apps for students to use. The interrater reliability of our MARS and MARuL scores was moderate, which was slightly lower than that described in the MARS and similar to the MARuL development. Although higher reliability might change scoring somewhat, it is unlikely to change our findings, as each individual reviewer identified the same top 3 apps, albeit in a different order, for both MARS and MARuL.

### Next Steps

The results across the 2 measures of app quality indicate the potential convergent and discriminant validity of raters' perceptions across MARS and MARuL. Further research to develop the construct validity of these 2 measures by using student outcome data with regard to highly rated apps will help to confirm their usefulness in their respective areas of focus.

### Conclusions

This systematic search for and evaluation of clinical skills mobile apps for perceived general quality and value for learning has highlighted the importance of using a fit-for-purpose measure of quality or value of mobile apps. The findings suggest that both MARS and MARuL instruments are useful and somewhat complementary. This study also highlights the variable quality of health-related education apps, likely because of the lack of regulation of health apps, in the iOS App Store and Google Play Store. However, Geeky Medics-OSCE revision and OSCE PASS are examples of how good practice in the development of apps can lead to quality apps for learning.

---

### Acknowledgments

This work was supported by a grant from the Otago Medical School Medical Education Research Fund. The authors would like to thank Michael Fauchelle from the University of Otago Wellington Medical and Health Sciences Library for his assistance with the initial literature search.

---

### Authors' Contributions

TG developed the initial research question and methodology, completed the app store searches, analyzed the data, and drafted the manuscript. GT assisted with the development of the methodology, reviewed the literature, completed the app store searches, assisted with writing the first draft, and edited the manuscript. SG assisted with the development of the methodology, reviewed the apps using MARS and MARuL, and edited the manuscript. JM reviewed the apps using MARS and MARuL and edited the manuscript. RG assisted with the development of the methodology and edited the manuscript.

---

## Conflicts of Interest

None declared.

## References

1. Mickan S, Tilson JK, Atherton H, Roberts NW, Heneghan C. Evidence of effectiveness of health care professionals using handheld computers: a scoping review of systematic reviews. *J Med Internet Res* 2013 Oct 28;15(10):e212 [FREE Full text] [doi: [10.2196/jmir.2530](https://doi.org/10.2196/jmir.2530)] [Medline: [24165786](https://pubmed.ncbi.nlm.nih.gov/24165786/)]
2. Paglialonga A, Lugo A, Santoro E. An overview on the emerging area of identification, characterization, and assessment of health apps. *J Biomed Inform* 2018 Jul;83:97-102 [FREE Full text] [doi: [10.1016/j.jbi.2018.05.017](https://doi.org/10.1016/j.jbi.2018.05.017)] [Medline: [29852318](https://pubmed.ncbi.nlm.nih.gov/29852318/)]
3. Scott IA, Scuffham P, Gupta D, Harch TM, Borch J, Richards B. Going digital: a narrative overview of the effects, quality and utility of mobile apps in chronic disease self-management. *Aust Health Rev* 2018 Nov;44(1):62-82. [doi: [10.1071/AH18064](https://doi.org/10.1071/AH18064)] [Medline: [30419185](https://pubmed.ncbi.nlm.nih.gov/30419185/)]
4. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
5. Sucala M, Cuijpers P, Muench F, Cardo R, Soflaur R, Dobrean A, et al. Anxiety: there is an app for that. A systematic review of anxiety apps. *Depress Anxiety* 2017 Jun;34(6):518-525. [doi: [10.1002/da.22654](https://doi.org/10.1002/da.22654)] [Medline: [28504859](https://pubmed.ncbi.nlm.nih.gov/28504859/)]
6. Meghani SH, MacKenzie MA, Morgan B, Kang Y, Wasim A, Sayani S. Clinician-targeted mobile apps in palliative care: a systematic review. *J Palliat Med* 2017 Oct;20(10):1139-1147. [doi: [10.1089/jpm.2017.0070](https://doi.org/10.1089/jpm.2017.0070)] [Medline: [28557549](https://pubmed.ncbi.nlm.nih.gov/28557549/)]
7. 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](https://doi.org/10.2196/mhealth.8613)] [Medline: [29592848](https://pubmed.ncbi.nlm.nih.gov/29592848/)]
8. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. *J Med Internet Res* 2015 Aug 17;17(8):e198 [FREE Full text] [doi: [10.2196/jmir.4581](https://doi.org/10.2196/jmir.4581)] [Medline: [26283290](https://pubmed.ncbi.nlm.nih.gov/26283290/)]
9. Alessa T, Abdi S, Hawley MS, de Witte L. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, usability, and user satisfaction. *JMIR Mhealth Uhealth* 2018 Jul 23;6(7):e10723 [FREE Full text] [doi: [10.2196/10723](https://doi.org/10.2196/10723)] [Medline: [30037787](https://pubmed.ncbi.nlm.nih.gov/30037787/)]
10. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for people with rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. *JMIR Mhealth Uhealth* 2017 Mar 21;5(2):e7. [doi: [10.2196/mhealth.6956](https://doi.org/10.2196/mhealth.6956)] [Medline: [28223263](https://pubmed.ncbi.nlm.nih.gov/28223263/)]
11. 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](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
12. Jonas-Dwyer DR, Clark C, Celenza A, Siddiqui ZS. Evaluating apps for learning and teaching. *Int J Emerg Technol Learn* 2012 Feb 29;7(1):54-57. [doi: [10.3991/ijet.v7i1.1901](https://doi.org/10.3991/ijet.v7i1.1901)]
13. Chuchu T, Nodoro T. An examination of the determinants of the adoption of mobile applications as learning tools for higher education students. *Int J Interact Mob Technol* 2019 Mar 25;13(03):53-67. [doi: [10.3991/ijim.v13i03.10195](https://doi.org/10.3991/ijim.v13i03.10195)]
14. Liaw SS, Huang HM. Perceived satisfaction, perceived usefulness and interactive learning environments as predictors to self-regulation in e-learning environments. *Comput Educ* 2013 Jan;60(1):14-24. [doi: [10.1016/j.compedu.2012.07.015](https://doi.org/10.1016/j.compedu.2012.07.015)]
15. Masters K, Ellaway RH, Topps D, Archibald D, Hogue RJ. Mobile technologies in medical education: AMEE Guide No. 105. *Med Teach* 2016 Jun;38(6):537-549. [doi: [10.3109/0142159X.2016.1141190](https://doi.org/10.3109/0142159X.2016.1141190)] [Medline: [27010681](https://pubmed.ncbi.nlm.nih.gov/27010681/)]
16. Boruff JT, Storie D. Mobile devices in medicine: a survey of how medical students, residents, and faculty use smartphones and other mobile devices to find information. *J Med Libr Assoc* 2014 Jan;102(1):22-30 [FREE Full text] [doi: [10.3163/1536-5050.102.1.006](https://doi.org/10.3163/1536-5050.102.1.006)] [Medline: [24415916](https://pubmed.ncbi.nlm.nih.gov/24415916/)]
17. Payne KF, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Med Inform Decis Mak* 2012;12:121 [FREE Full text] [doi: [10.1186/1472-6947-12-121](https://doi.org/10.1186/1472-6947-12-121)] [Medline: [23110712](https://pubmed.ncbi.nlm.nih.gov/23110712/)]
18. Snashall E, Hindocha S. The use of smartphone applications in medical education. *Open Medicine J* 2016 Dec 27;3(1):322-327. [doi: [10.2174/1874220301603010322](https://doi.org/10.2174/1874220301603010322)]
19. Wallace S, Clark M, White J. 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. *BMJ Open* 2012;2(4):e001099 [FREE Full text] [doi: [10.1136/bmjopen-2012-001099](https://doi.org/10.1136/bmjopen-2012-001099)] [Medline: [22923627](https://pubmed.ncbi.nlm.nih.gov/22923627/)]
20. Mackay BJ, Anderson J, Harding T. Mobile technology in clinical teaching. *Nurse Educ Pract* 2017 Jan;22:1-6. [doi: [10.1016/j.nepr.2016.11.001](https://doi.org/10.1016/j.nepr.2016.11.001)] [Medline: [27871040](https://pubmed.ncbi.nlm.nih.gov/27871040/)]
21. Dimond R, Bullock A, Lovatt J, Stacey M. Mobile learning devices in the workplace: 'as much a part of the junior doctors' kit as a stethoscope'? *BMC Med Educ* 2016 Aug 17;16(1):207 [FREE Full text] [doi: [10.1186/s12909-016-0732-z](https://doi.org/10.1186/s12909-016-0732-z)] [Medline: [27530343](https://pubmed.ncbi.nlm.nih.gov/27530343/)]
22. Harden RM, Lilley P. *The Eight Roles of the Medical Teacher*. Edinburgh: Elsevier; 2018:1-328.



23. Thistlethwaite JE, Davies D, Ekeocha S, Kidd JM, MacDougall C, Matthews P, et al. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. *Med Teach* 2012;34(6):421-444. [doi: [10.3109/0142159X.2012.680939](https://doi.org/10.3109/0142159X.2012.680939)] [Medline: [22578051](https://pubmed.ncbi.nlm.nih.gov/22578051/)]
24. Lee JS, Kim SW. Validation of a tool evaluating educational apps for smart education. *J Educ Comput Res* 2015 Mar 26;52(3):435-450. [doi: [10.1177/0735633115571923](https://doi.org/10.1177/0735633115571923)]
25. Gladman T, Tylee G, Gallagher S, Mair J, Rennie S, Grainger R. A tool for rating the value of health education mobile apps to enhance student learning (MARuL): development and usability study. *JMIR Mhealth Uhealth* 2020 Jul 31;8(7):e18015 [FREE Full text] [doi: [10.2196/18015](https://doi.org/10.2196/18015)] [Medline: [32735228](https://pubmed.ncbi.nlm.nih.gov/32735228/)]
26. Snow CE, Torous J, Gordon-Elliott JS, Penzner JB, Meyer F, Boland R. Use of electronic resources for psychiatry clerkship learning: a medical student survey. *Acad Psychiatry* 2017 Oct;41(5):656-660. [doi: [10.1007/s40596-016-0647-3](https://doi.org/10.1007/s40596-016-0647-3)] [Medline: [27995539](https://pubmed.ncbi.nlm.nih.gov/27995539/)]
27. Task Force on the Clinical Skills Education of Medical Students. Recommendations for clinical skills curricula for undergraduate medical education. Association of American Medical Colleges. 2005. URL: [https://store.aamc.org/downloadable/download/sample/sample\\_id/174/](https://store.aamc.org/downloadable/download/sample/sample_id/174/) [accessed 2021-03-30]
28. 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 Dec;26(4):2930-2945 [FREE Full text] [doi: [10.1177/1460458220952917](https://doi.org/10.1177/1460458220952917)] [Medline: [32914696](https://pubmed.ncbi.nlm.nih.gov/32914696/)]
29. Do we really need App logins? Apptimize. 2015. URL: <https://apptimize.com/blog/2015/04/do-we-really-need-app-logins/> [accessed 2021-02-01]
30. Brandall B. First-time use: how to reduce the initial friction of app usage. Usability Geek. URL: <https://usabilitygeek.com/first-time-use-how-to-reduce-initial-friction-of-app-usage/> [accessed 2021-02-01]
31. Liu H, Ofstad A, Nicholas E. Airtable. 2019. URL: <https://airtable.com/> [accessed 2019-03-30]
32. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Br Med J* 2009 Jul 21;339:b2535 [FREE Full text] [doi: [10.1136/bmj.b2535](https://doi.org/10.1136/bmj.b2535)] [Medline: [19622551](https://pubmed.ncbi.nlm.nih.gov/19622551/)]
33. RStudio Team. RStudio: integrated development environment for R. 2019. URL: <https://www.rstudio.com/products/rstudio/> [accessed 2020-03-30]
34. 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](https://doi.org/10.1016/j.jcm.2016.02.012)] [Medline: [27330520](https://pubmed.ncbi.nlm.nih.gov/27330520/)]
35. Mack J, Mack R. Fnd: experience the App store and iTunes anywhere. 2021. URL: <https://fnd.io/> [accessed 2021-01-27]
36. Awan O, Dey C, Salts H, Brian J, Fotos J, Royston E, et al. Making learning fun: gaming in radiology education. *Acad Radiol* 2019 Aug;26(8):1127-1136. [doi: [10.1016/j.acra.2019.02.020](https://doi.org/10.1016/j.acra.2019.02.020)] [Medline: [31005406](https://pubmed.ncbi.nlm.nih.gov/31005406/)]
37. Rotgans JI, Schmidt HG. The role of interest in learning: knowledge acquisition at the intersection of situational and individual interest. In: O'Keefe P, Harackiewicz J, editors. *The Science of Interest*. Switzerland: Springer; 2017:69-93.
38. Thoman DB, Sansone C, Geerling D. The dynamic nature of interest: embedding interest within self-regulation. In: O'Keefe P, Harackiewicz J, editors. *The Science of Interest*. Switzerland: Springer; 2017:27-47.
39. Hurst EJ. Evaluating health and wellness mobile applications. *J Hosp Librariansh* 2018 Jul 17;18(3):266-278. [doi: [10.1080/15323269.2018.1472928](https://doi.org/10.1080/15323269.2018.1472928)]
40. Böhme C, von Osthoff MB, Frey K, Hübner J. Development of a rating tool for mobile cancer apps: information analysis and formal and content-related evaluation of selected cancer apps. *J Cancer Educ* 2019 Feb;34(1):105-110. [doi: [10.1007/s13187-017-1273-9](https://doi.org/10.1007/s13187-017-1273-9)] [Medline: [28819928](https://pubmed.ncbi.nlm.nih.gov/28819928/)]
41. Fiore P. How to evaluate mobile health applications: a scoping review. *Stud Health Technol Inform* 2017;234:109-114. [Medline: [28186025](https://pubmed.ncbi.nlm.nih.gov/28186025/)]
42. 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](https://doi.org/10.2196/13080)] [Medline: [30747715](https://pubmed.ncbi.nlm.nih.gov/30747715/)]
43. Jamaladin H, van de Belt TH, Luijpers LC, de Graaff FR, Bredie SJ, Roeleveld N, et al. Mobile apps for blood pressure monitoring: systematic search in app stores and content analysis. *JMIR Mhealth Uhealth* 2018 Nov 14;6(11):e187 [FREE Full text] [doi: [10.2196/mhealth.9888](https://doi.org/10.2196/mhealth.9888)] [Medline: [30429116](https://pubmed.ncbi.nlm.nih.gov/30429116/)]
44. LeBeau K, Huey LG, Hart M. Assessing the quality of mobile apps used by occupational therapists: evaluation using the user version of the mobile application rating scale. *JMIR Mhealth Uhealth* 2019 May 01;7(5):e13019 [FREE Full text] [doi: [10.2196/13019](https://doi.org/10.2196/13019)] [Medline: [31066712](https://pubmed.ncbi.nlm.nih.gov/31066712/)]
45. Virani A, Duffett-Leger L, Letourneau N. Parenting apps review: in search of good quality apps. *Mhealth* 2019;5:44 [FREE Full text] [doi: [10.21037/mhealth.2019.08.10](https://doi.org/10.21037/mhealth.2019.08.10)] [Medline: [31620471](https://pubmed.ncbi.nlm.nih.gov/31620471/)]

## Abbreviations

- ICC:** intraclass correlation coefficient  
**MARS:** Mobile App Rating Scale

**MARuL:** Mobile App Rubric for Learning

**OSCE:** objective structured clinical examination

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

*Edited by L Buis; submitted 29.10.20; peer-reviewed by N Moore, M Fajardo; comments to author 15.12.20; revised version received 01.02.21; accepted 18.03.21; published 23.04.21.*

*Please cite as:*

*Gladman T, Tylee G, Gallagher S, Mair J, Grainger R*

*Measuring the Quality of Clinical Skills Mobile Apps for Student Learning: Systematic Search, Analysis, and Comparison of Two Measurement Scales*

*JMIR Mhealth Uhealth 2021;9(4):e25377*

*URL: <https://mhealth.jmir.org/2021/4/e25377>*

*doi: [10.2196/25377](https://doi.org/10.2196/25377)*

*PMID: [33890859](https://pubmed.ncbi.nlm.nih.gov/33890859/)*

©Tehmina Gladman, Grace Tylee, Steve Gallagher, Jonathan Mair, Rebecca Grainger. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 23.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Effects of App-Based Transitional Care on the Self-Efficacy and Quality of Life of Patients With Spinal Cord Injury in China: Randomized Controlled Trial

Ting Liu<sup>1\*</sup>, MSN; Sumei Xie<sup>2\*</sup>, BSN; Yingmin Wang<sup>3</sup>, MSN; Jie Tang<sup>4</sup>, BSN; Xiaokuo He<sup>5</sup>, PhD; Tiebin Yan<sup>3</sup>, PhD; Kun Li<sup>6</sup>, PhD

<sup>1</sup>Department of Rehabilitation Medicine, The Eighth Affiliated Hospital, Sun Yat-sen University, Shenzhen, China

<sup>2</sup>Department of Spinal Cord Injury Rehabilitation, Guangdong Provincial Work Injury Rehabilitation Hospital, Guangzhou, China

<sup>3</sup>Department of Rehabilitation Medicine, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

<sup>4</sup>Department of Spinal Cord Injury Rehabilitation, Sichuan Provincial Rehabilitation Hospital, Chengdu, China

<sup>5</sup>Department of Rehabilitation Medicine, The Fifth Hospital of Xiamen, Xiamen, China

<sup>6</sup>School of Nursing, Sun Yat-sen University, Guangzhou, China

\*these authors contributed equally

**Corresponding Author:**

Kun Li, PhD

School of Nursing

Sun Yat-sen University

No 74 Zhong Shan Second Road

Guangzhou, 510080

China

Phone: 86 13822206519

Email: [likun22@mail.sysu.edu.cn](mailto:likun22@mail.sysu.edu.cn)

## Abstract

**Background:** Spinal cord injury (SCI) severely impairs the physical and mental health of patients, decreasing their self-efficacy in coping with daily life and quality of life (QOL). In China, a large gap remains between the complex long-term health needs of SCI patients and the current community care system. With the prevalence of mobile terminals, the usage of mobile health apps has the potential to fill this gap by extending qualified medical resources to the families of SCI patients. Our team developed the app *Together* for the transitional care of home-dwelling SCI patients in China.

**Objective:** This study aimed to evaluate the effects of app-based transitional care on the self-efficacy and QOL of SCI patients.

**Methods:** Through a three-round Delphi process, an Android app was designed. Both medical staff and patients could access the app. Medical staff used it for providing remote transitional care to SCI patients. Patients used it to view transitional care time and send messages to medical staff. Thereafter, a multicenter and assessor-blinded randomized controlled trial was conducted. Participants (n=98) who had SCI and lived at home following discharge were recruited and randomly assigned to a study group (n=49) and control group (n=49) using a randomized number list in four research centers. Patients in both groups received systematic discharge education before discharge. The study group received five follow-ups conducted by trained nurses through the app, which had four core functions, namely remote assessment, health education, interdisciplinary referral, and patient interaction, at weeks 2, 4, 6, 8, and 12 following discharge. The control group received a routine telephone follow-up conducted by nurses at week 12 following discharge. The outcome measures were the Moorong Self-Efficacy Scale (MSES) and 36-item Short-Form Health Survey (SF-36) scores. Data were collected before discharge (T<sub>0</sub>) and at weeks 12 (T<sub>1</sub>) and 24 following discharge (T<sub>2</sub>). Differences between the groups were tested by repeated measures analysis of variance and simple effect analysis.

**Results:** After the follow-up, the total MSES scores in the study group improved over time (T<sub>0</sub>=67.80, T<sub>1</sub>=71.90, and T<sub>2</sub>=76.29) and were higher than those in the control group (T<sub>2</sub>=64.49) at 24 weeks following discharge (simple effect analysis:  $F_1=8.506$ ,  $P=.004$ ). Regarding the total SF-36 score, although it was higher in patients from the study group (T<sub>2</sub>=65.36) than those from the control group (T<sub>2</sub>=58.77) at 24 weeks following discharge, only time effects were significant ( $F_{2,95}=6.671$ ,  $P=.002$ ) and neither

the group effects nor the interaction effects influenced the change in QOL (group effects:  $F_{1,96}=0.082$ ,  $P=.78$ ; interaction effects:  $F_{2,95}=3.059$ ,  $P=.052$ ).

**Conclusions:** This study confirmed that app-based transitional care improves the self-efficacy of SCI patients. Nevertheless, QOL improvement is not yet evident. Future investigations with larger sample sizes and longer observation periods are warranted to further verify the effects.

**Trial Registration:** Chinese Clinical Trial Registry ChiCTR-IPR-17012317; <http://www.chictr.org.cn/showproj.aspx?proj=19828>

(*JMIR Mhealth Uhealth* 2021;9(4):e22960) doi:[10.2196/22960](https://doi.org/10.2196/22960)

## KEYWORDS

spinal cord injury; mobile app; transitional care; self-efficacy; quality of life

## Introduction

Spinal cord injury (SCI) is often a serious and life-changing disease with approximately 200,000 to 500,000 individuals newly diagnosed annually worldwide [1,2]. In China, the annual incidence of SCI is 23.7 to 60.6 persons per million [3-5]. SCI is a disastrous event for patients and their families owing to the associated lifelong disabilities and types of complications, which lead to deterioration of functioning independence [6], higher depression [7], higher readmission [8], poorer self-efficacy [1], and more problems compared with those observed in individuals without SCI. It can decrease one's quality of life (QOL) and have high care and treatment burdens.

In China, most newly injured patients with SCI live at home after acute treatment and subacute rehabilitation [9]. They urgently require professional medical support and rehabilitation services in communities. However, there is a gap between the health needs of patients with SCI and the existing developing community care system in China. The complexity of SCI means there are more professional requirements for community medical care. The present community medical resources across China remain imbalanced. Many communities are unable to meet the needs of patients with SCI based on the competency of health professionals and basic community medical facilities [10]. Some well-equipped comprehensive hospitals and rehabilitation centers have attempted to fill the gap through transitional care, which extends professional medical care to the families of patients with SCI. In China, for patients with SCI who have just returned home from hospitals, transitional care involves continuing care services, and medical staff members provide patients with self-management knowledge and skills to help them avoid complications and unplanned readmissions, and better adapt to family and social life. The most common formats are home visits, telephone follow-ups, and outpatient services. However, some limitations continue to exist in these approaches, such as a small number of beneficiaries because of distance and insurance [11,12] and inconvenience for patients with motor disabilities receiving outpatient services [13].

With the development of the internet and prevalence of mobile terminals, such as intelligent cell phones and tablets, mobile health is playing a more important role in the Chinese medical service. A mobile health app is a form of mobile health that is installed on mobile electronic devices and has a variety of functions, including assisting in diagnosis, tele-health education, and follow-up [14]. Nowadays, apps are widely applied in the

transitional care of patients with chronic diseases [15], postoperative patients [16,17], and puerperium women [18]. Numerous studies have demonstrated that app-based transitional care could enhance self-efficacy [15,18,19], prevent complications [16,17], improve QOL [20], and reduce readmissions [21] among patients.

With the development of modern medicine, one of the ultimate goals of rehabilitation for patients with SCI is to improve their QOL. QOL has been identified as one of the most common outcome indicators for patients with SCI, which could provide a comprehensive picture of the individual's physical-psychological health and social domains [22]. Owing to the different degrees of dysfunction, patients with SCI have a marked decline in QOL compared with the general population [23]. An improvement in QOL suggests an adaptive process operating over a long period of time [24]. Interestingly, self-efficacy has been shown to be an important predictor of QOL in patients with SCI [22]. Self-efficacy is a psychological concept that describes the confidence of a person in performing specific activities and pursuing desired goals [25]. For patients with SCI, self-efficacy includes general self-efficacy and SCI-specific self-efficacy, such as self-efficacy in coping with daily life and in self-management [26]. SCI patients with low levels of self-efficacy tend to have a poor QOL [22]. In addition, low self-efficacy is strongly associated with depressive moods in patients with SCI [27]. The improvement of self-efficacy in patients with SCI is likely to promote life-long behaviors of daily life and self-regulation for health maintenance [28,29]. Therefore, improvement of the QOL and self-efficacy of patients with SCI was the focus of this study.

In view of the superiority of apps, our research team developed an app, named *Together*, specifically for the transitional care of patients with SCI. At present, our app has two user interfaces (medical staff and patients). The current language is Chinese, and the copyright is held by Sun Yat-sen University [30]. The app *Together* includes the following four core functions: remote assessment, health education, interdisciplinary referral, and patient interaction. The aim of this study was to test the effects of app-based transitional care on the self-efficacy and QOL of patients with SCI through a randomized controlled trial (RCT), thereby providing clinical evidence and guidance for the development of transitional care for patients with SCI in China.

## Methods

### Study Design and Settings

This research was a multicenter and assessor-blinded RCT (Chinese Clinical Trial Registry, ChiCTR-IPR-17012317). It was implemented in four research centers, which included one rehabilitation department in a comprehensive hospital and three rehabilitation hospitals in Guangzhou, Chengdu, and Shiyuan in China. We conducted RCTs in each of the research centers over the same period.

### Development of the App

The app *Together* was developed based on the International Classification of Functioning, Disability, and Health (ICF), which offered a platform for health professionals to provide transitional care for patients with SCI at home in China. The development process of the app is scientific and rigorous. A three-round Delphi expert panel and expert consultation were performed step by step to build the framework of the app, which included the contents of remote transitional care, including SCI follow-up indicators, measurement methods, and standardized guidelines [31,32].

With regard to the construction of the SCI follow-up indicators, an ICF set for the transitional care outcome indicators of patients with SCI in China was developed through a three-round Delphi expert panel in preliminary studies [31,32]. ICF is a unified and standard terminology system for multidisciplinary use. It provides a comprehensive perspective in describing one's functioning, and the interdisciplinary focus rendered it more suitable for multidisciplinary teamwork [33]. Each ICF category can reflect the individual's function by using ICF qualifiers (divided into 0 to 4 levels). In our study, ICF categories were the follow-up indicators to evaluate SCI patients' functioning. A total of 52 experts took part in the Delphi process, all of whom were nursing experts and were certified rehabilitation nurses with at least 5 years of clinical experience in nursing. Considering the feasibility of clinical practice, the outcome indicators suitable for SCI follow-up (32 for large follow-up and 12 for small follow-up) (Table 1) were selected from the ICF set above by an expert panel consisting of two rehabilitation physicians and three nursing specialists in SCI. In this study, the medical staff provided the "small follow-up" intervention at 2, 4, 6, and 8 weeks following discharge, with a total of 12 follow-up indicators. It was considered that patients needed frequent follow-ups within 3 months when they returned home, so we set less follow-up indicators to meet clinical needs. The medical staff provided the "large follow-up" intervention at 12 weeks following discharge, with a total of 32 follow-up indicators. Since patients with SCI had been discharged from the hospital for a period of time, a regular systematic follow-up for them was necessary, so we set more large follow-up

indicators to comprehensively assess the patient's current functioning. To test the validity of the SCI follow-up indicators, we used Rasch analysis [34] to examine each component of the ICF set and the overall fit to a Rasch model. The results [35] of the Rasch analysis showed good fit to the Rasch model for the different components of the ICF set in the app after modification. Both overall and single-item fits were satisfactory. These results indicated the suitability of our selected ICF set as follow-up indicators to evaluate the functioning of patients with SCI.

With regard to the construction of measurement methods and standardized guidelines, for each follow-up indicator, the operational measurement method was set based on literature review and existing measurements by expert consultation [36]. We transformed patient information into ICF qualifiers according to different criteria, such as percentages, frequency, and medical staff assessment results. For example, for *b4200 Increased blood pressure* and *b4201 Decreased blood pressure*, we transformed patient information into ICF qualifiers according to frequency. Stable blood pressure over the past month was rated as 0, whereas high/low blood pressure almost every day received a rating of 4. Our app can automatically transform routine clinical assessments into the unified ICF qualifiers ("0," no problem; "1," mild problem; "2," moderate problem; "3," serious problem; "4," complete problem; and "9," not applicable) for better understanding among the multidisciplinary team. Moreover, a standardized health education program for each follow-up indicator was developed by expert consultation based on the Knowledge-Attitude-Practice theory [37].

Based on previous work, software engineers developed the mobile app using the Java language according to the requirements set by the researchers [31]. Eventually, an Android version of the app was designed. Visitors to the app could be medical staff or patients, but they were required to apply for registration with their real identity first. Thereafter, the researchers verified applicants' information before they could log in. Unregistered users could not access the app even if they had downloaded it. In addition, medical staff and patients had different access interfaces after logging in. Medical staff entered a follow-up system and used it to provide transitional care for patients with SCI. The patient interface could be only used to view follow-up time and send messages to the medical staff. For each follow-up, medical staff needed to record the patient's follow-up outcomes in the app. The collected information was stored in the cloud, which was actually a cloud computing server like Google cloud for data storage with good security. The security measures of our app were as follows: (1) user login required a password for secure login; (2) users' sensitive information was encrypted and saved to prevent data leakage; and (3) data were regularly saved to the cloud to prevent loss.

**Table 1.** Follow-up outcome indicators of transitional care for patients with spinal cord injury.

Classification	Variable	Small follow-up	Large follow-up <sup>a</sup>
ICF <sup>b</sup> category	Body function	<i>b280 Sense of pain</i> <i>b415 Blood vessel functions</i> <i>b435 Immunological system functions</i> <i>b440 Functions of the respiratory system</i> <i>b4200 Increased blood pressure</i> <i>b4201 Decreased blood pressure</i> <i>b530 Weight maintenance functions</i>	<i>b134 Sleep functions</i> <i>b152 Emotional functions</i> <i>b280 Sense of pain</i> <i>b415 Blood vessel functions</i> <i>b4200 Increased blood pressure</i> <i>b4201 Decreased blood pressure</i> <i>b435 Immunological system functions</i> <i>b440 Functions of the respiratory system</i> <i>b530 Weight maintenance functions</i> <i>b640 Sexual functions</i> <i>b660 Procreation functions</i> <i>b710 Mobility of joint functions</i> <i>b730 Muscle power functions</i> <i>b735 Muscle tone functions</i>
ICF category	Body structure	<i>s810 Structure of areas of skin</i>	<i>s810 Structure of areas of skin</i>
ICF category	Activities and participation	<i>d5300 Regulating urination</i> <i>d5301 Regulating defecation</i> <i>d410 Changing basic body position</i> <i>d420 Transferring oneself</i>	<i>d410 Changing basic body position</i> <i>d420 Transferring oneself</i> <i>d450 Walking</i> <i>d465 Moving around using equipment</i> <i>d510 Washing oneself</i> <i>d520 Caring for body parts</i> <i>d5300 Regulating urination</i> <i>d5301 Regulating defecation</i> <i>d540 Dressing</i> <i>d550 Eating</i> <i>d560 Drinking</i> <i>d9205 Socializing</i> <i>d760 Family relationships</i>
ICF category	Contextual factors <sup>c</sup>	N/A <sup>d</sup>	<i>e1201 Assistive products and technology for personal indoor and outdoor mobility and transportation</i> <i>e155 Design, construction, and building products and technology of buildings for private use</i> <i>Acceptance of life in a wheelchair/in bed</i> <i>Coping with everyday life</i> <i>Adjustment to new body image</i> <i>Knowledge about spinal cord injury</i>
Concepts not covered in the ICF <sup>e</sup>	N/A	<i>Ability of the caregiver</i>	<i>Ability of the caregiver</i>

<sup>a</sup>The large follow-up indicators include all small follow-up indicators.

<sup>b</sup>ICF: International Classification of Functioning, Disability, and Health.

<sup>c</sup>The contextual factors consisted of environmental factors and personal factors. Considering medical staff's heavy workload, it is not realistic to evaluate four components of ICF at each follow-up. The small follow-up was conducted within 3 months following discharge. During this time, we paid more attention to the health conditions of patients. Thus, the contextual factors were not included in the small follow-up.

<sup>d</sup>N/A: not applicable.

<sup>e</sup>Through the Delphi process, we exacted one concept not covered in the ICF (ability of the caregiver). This reflected the views of Chinese experts.

## Participants

Participants were recruited from May 2018 to December 2019. The inclusion criteria were as follows: (1) diagnosis of complete or incomplete SCI according to the International Standards for the Neurological Classification of SCI of the American Spinal

Injury Association, 2016 [38], with confirmation by computed tomography/magnetic resonance imaging (either traumatic or nontraumatic); (2) onset of SCI within the past 2 years; (3) being alert and conscious, and capable of daily verbal communication; (4) age >18 years; (5) living at home following discharge; (6) availability of mobile terminals and internet access at home;

and (7) familiarity with the usage of mobile terminals. The exclusion criteria were as follows: (1) congenital spinal cord disease; (2) severe cardiovascular, brain, pulmonary, liver, or renal complications; and (3) admission to other medical institutions following discharge.

Randomization was performed using the software Research Randomizer [39]. This software generated a random number list for each center with two numbers in the list (“1” and “2”), which referred to the study group and control group, respectively. According to the discharge sequence and corresponding randomized number, the recruited patients were randomly assigned to the study group and control group. The unused numbers were discarded after completion of the study.

The sample size was estimated based on the primary outcome measure (Moorong Self-Efficacy Scale [MSES] score) between the groups. In the study conducted by Martin et al [40], the mean difference in the MSES score at week 12 was normally distributed, with a mean of 1.95 and a standard deviation (SD) of 2.77. A total of 84 patients (42 patients in each group) were required for the study, with a two-tailed  $\alpha$  of .05 and a power of 0.80. We planned to recruit 98 patients in total to allow for a 15% dropout rate.

### Pilot Study

Prior to the RCT, a pilot study was conducted. First, for testing the functions of the app, we selected a multidisciplinary team that included two nurses, a rehabilitation physician, a physical therapist, and an occupational therapist, and 20 patients with SCI who met the inclusion criteria to download and use the app at one research center in Guangzhou. Owing to limited time, we only tested the first small follow-up. Both medical staff and patients were asked about the app’s user experience through a face-to-face interview. From the interview, we learned that medical staff generally believed that the app can help them to carry out follow-up visits in terms of comprehensive content, a user-friendly interface, and simple operation, and patients found it easy to contact their health care providers through the app. Subsequently, the researchers and software engineers optimized the app according to the suggestions. Second, for testing the validity and reliability of the outcome measures, 20 patients were asked to fill out the MSES and QOL scale at baseline.

### Interventions

Both the study and control groups received routine systematic discharge education provided by the medical team in the research centers before discharge, including information on medication, bladder management, bowel management, respiratory training, activities of daily living training, and prevention of complications (pressure ulcers, infections, falls, autonomic dysfunction, orthostatic hypotension, deep vein thrombosis, etc).

### Study Group

A multidisciplinary team, including one follow-up nurse, one rehabilitation physician, one physical therapist, and one occupational therapist, managed the app-based transitional care in each center. The nurse was in charge of the team and responsible for the app-based transitional care. All follow-up nurses were required to have over 3 years of working experience in caring for patients with SCI.

The follow-up nurses provided information on the installation and usage of the app to the patients and their main caregivers in the study group prior to discharge. Five follow-ups based on the app were conducted for the study group by the multidisciplinary team at 2, 4, 6, 8, and 12 weeks following discharge. The timing and frequency of postdischarge interventions were set according to previous studies and the feasibility of this study. A longitudinal study suggested that the QOL of patients with SCI at 3 months following discharge was an important predictor of QOL at 15 months [41]. One telephone counselling program that delivered seven tele-counseling sessions over a 12-week period was effective in managing psychological outcomes [42]. Therefore, the researchers focused on patients with SCI who were discharged from the rehabilitation center within 3 months and developed the transitional care plan with five follow-ups based on the app within 3 months following discharge. The app would automatically remind the follow-up nurse of the patients that had to be contacted within 1 week prior to the end of the follow-up. The app-based transitional care intervention provided by the multidisciplinary team for patients with SCI included the following four parts: (1) remote health assessment, (2) health education, (3) interdisciplinary referral, and (4) patient interaction.

### Remote Health Assessment

In the follow-up system of the medical staff interface, ICF categories were displayed as follow-up indicators. The nurse could remotely evaluate patients’ functioning by questioning patients via telephone according to preset instructions on the app. A verbal prompt was provided in the medical staff interface of the app to guide them in implementing a standard assessment, and the app automatically transformed the original clinical assessment results to the ICF qualifiers according to preset operational measuring guidelines. For example, for *b280 Sense of pain* (Figure 1), the verbal prompt was “If 0 is not painful and 10 is the most painful, how serious is your pain?” After the nurse selected the appropriate number from 0 to 10 according to the patient’s answer, the app automatically transformed the assessment results to ICF qualifiers. If the patient’s condition did not change greatly from the last follow-up, the nurse could click on “The current condition is the same as the last follow-up.” Subsequently, the system could achieve this automatically. The setting can automatically synchronize the results of the last follow-up, which not only facilitates the medical staff to obtain the patients’ previous functional status, but also saves their time.

**Figure 1.** Remote assessment interface for the category of "b280 Sense of pain." ICF评分值 means "ICF qualifiers." After clicking, ICF qualifiers will be shown. 话术 means "verbal prompt." After clicking, preset instructions will be shown. 病情同上次随访 means "The current condition is the same as the last follow-up." 电话联系 means "Telephone." The phone number of the patient is shown here. After clicking, the follow-up nurse can contact the patient by telephone. ICF: International Classification of Functioning, Disability, and Health.

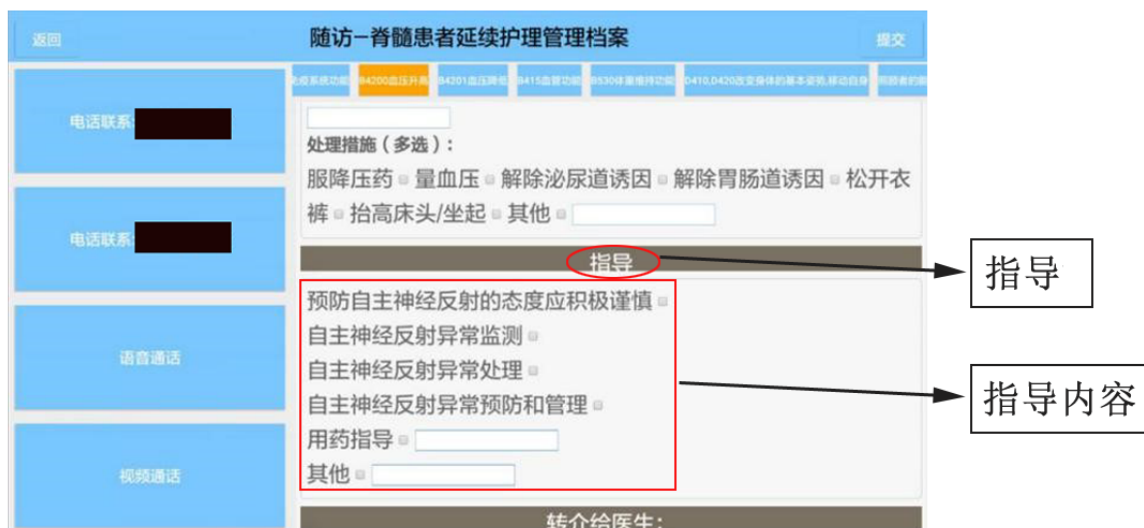


**Health Education**

The app provided a standard health education framework on each ICF category based on the Knowledge-Attitude-Practice theory. The health education framework of each ICF category was only visualized in the medical staff interface of the app. The follow-up nurses combined their own clinical experience and knowledge to provide health education for patients according to the prompted and standard health education

framework in the app. Figure 2 shows the health education framework for the ICF category of b4200 Increased blood pressure in the app. The education program included emphasizing the importance of preventing autonomic dysreflexia (AD), checking the monitoring records regarding AD, reviewing emergency management for the occurrence of AD, assessing the prevention and management of AD, providing medication guidance, etc.

**Figure 2.** Health education interface for the category of "b4200 Increased blood pressure." 指导 means "Health education." 指导内容 means "Health education contents." The health education contents for the ICF category "b4200 Increased blood pressure" are as follows: (1) the attitude toward preventing AD is positive and cautious; (2) instructing patients in the monitoring of AD; (3) instructing patients in the treatment of AD; (4) instructing patients in the prevention and management of AD; (5) medication guidance; and (6) others. AD: autonomic dysreflexia; ICF: International Classification of Functioning, Disability, and Health.



**Interdisciplinary Referral**

The referral function is a closed loop processing mode. When the follow-up nurses encountered problems that did not belong to the SCI nursing specialty, such as drug use and home modifications, referral messages were sent by the nurses to the

corresponding members of the multidisciplinary team (eg, rehabilitation physician, physical therapist, or occupational therapist) through the app. The team members were required to log in to the app within 3 days and handle the referral problems by contacting the patients. In the end, the physician or therapist



needed to fill in the referral record in the app. If the specialist was unable to contact the patient within 3 days, the referral record was not allowed to be filled in anymore and the patient

was forced to withdraw from our study. [Figure 3](#) shows an occupational therapist using the app.

**Figure 3.** An occupational therapist using the app.



### Patient Interaction

Medical staff could communicate with patients via telephone and the messaging function of the app. In turn, patients could contact medical staff by sending messages. Before every follow-up or contacting patients, medical staff read the messages received. In addition, nurses dealt with the messages from patients every Monday. If they encountered problems that they could not handle, they needed to send referral messages to other specialists. Relevant specialists were required to log in to the app to contact patients within 3 days.

### Control Group

No transitional care service was provided to the control group in the 3 months following discharge. At 12 weeks following discharge, the patients in the control group received a routine telephone follow-up conducted by nurses to mainly assess the functional level and complications of the patients and the disease management knowledge and skills of the patients and their caregivers. Corresponding health education was also included.

### Outcome Measures

Demographic and disease-related data were collected before discharge. All patients were assessed using the MSES and 36-item Short-Form Health Survey (SF-36) at baseline (prior to discharge;  $T_0$ ), 12 weeks following discharge ( $T_1$ ), and 24 weeks following discharge ( $T_2$ ).

### Demographic Disease Characteristics

The demographic disease inquiry consisted of two parts, namely demographic information and disease information. Patients were required to fill out the demographic information (name, gender, age, education background, marital status, insurance, occupation status, per capita income, main caregivers, etc), and nurses filled

out the disease information (diagnosis, injury level, injury severity, causes, disease duration, etc) according to medical records.

### MSES

MSES is a 16-item self-report scale developed by Middleton that measures one's ability to control behavior and outcomes, specifically among patients with SCI [43]. It consists of three-factor structures (social function self-efficacy, general self-efficacy, and personal function self-efficacy), and utilizes a 7-point Likert scale ranging from 1 (very uncertain) to 7 (very certain). Higher scores indicate better self-efficacy of patients. In a previous study, it exhibited adequate reliability (Cronbach  $\alpha$  of .94) and validity (content validity index [CVI] of 0.91) [44,45]. With the permission of the original authors, we developed a Chinese version of the MSES through forward and backward translation. In the study, the scale CVI was 0.99 (evaluated by eight clinical nursing specialists) and the item CVI ranged from 0.88 to 1.00. In addition, the Cronbach  $\alpha$  was .91 in a pilot test.

### SF-36

SF-36 is commonly used to measure the health-related QOL of patients with SCI [46]. It contains 36 items that measure perceived health in eight domains (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health), with higher scores (range 0-100) reflecting better perceived health. The first four domains form the physical component summary (PCS), and the last four domains form the mental component summary (MCS). The Chinese version of the SF-36 was translated in 1998 [47]. When applied in patients with SCI, the Cronbach  $\alpha$  values of the domains of the SF-36 ranged from .76 to .90 [48].

## Procedure

First, face-to-face training for the follow-up team members of each research center was conducted simultaneously. The training lasted 4 hours and consisted of an introductory lecture on the study and app, as well as a workshop on the operation of the app. All members participated in a qualification test after training, with a minimum requirement score of  $\geq 80/100$ .

Subsequently, the RCT was performed after the training. Eligible patients were recruited prior to discharge and randomly assigned to the study or control group according to their sequence of discharge. Both groups received routine systematic discharge education, and the patients in the study group received guidance on the usage of the app. The patients or medical staff downloaded the app by scanning a QR code. A verification system was set in the app, and only patients in the study group could pass the identity verification and use the patient interface of the app. The study group underwent five app-based follow-ups (at 2, 4, 6, 8, and 12 weeks following discharge) performed by the multidisciplinary team, while the control group underwent one routine telephone follow-up conducted by nurses. The outcome measures were collected by researchers at  $T_0$ ,  $T_1$ , and  $T_2$  by sending online survey links to the mobile phones of patients. After completing the data collection, patients received a US \$3 reward online.

## Ethical Considerations

The study protocol was approved by the ethics committee of Sun Yat-sen University (2017ZSLEYEC-062). All patients were informed of the research process, the aims, their privacy protection, and their rights. All patients were required to sign the informed consent form. For patients who could not sign the consent form owing to disabilities, consent was considered if the patient verbally agreed in the presence of a witness.

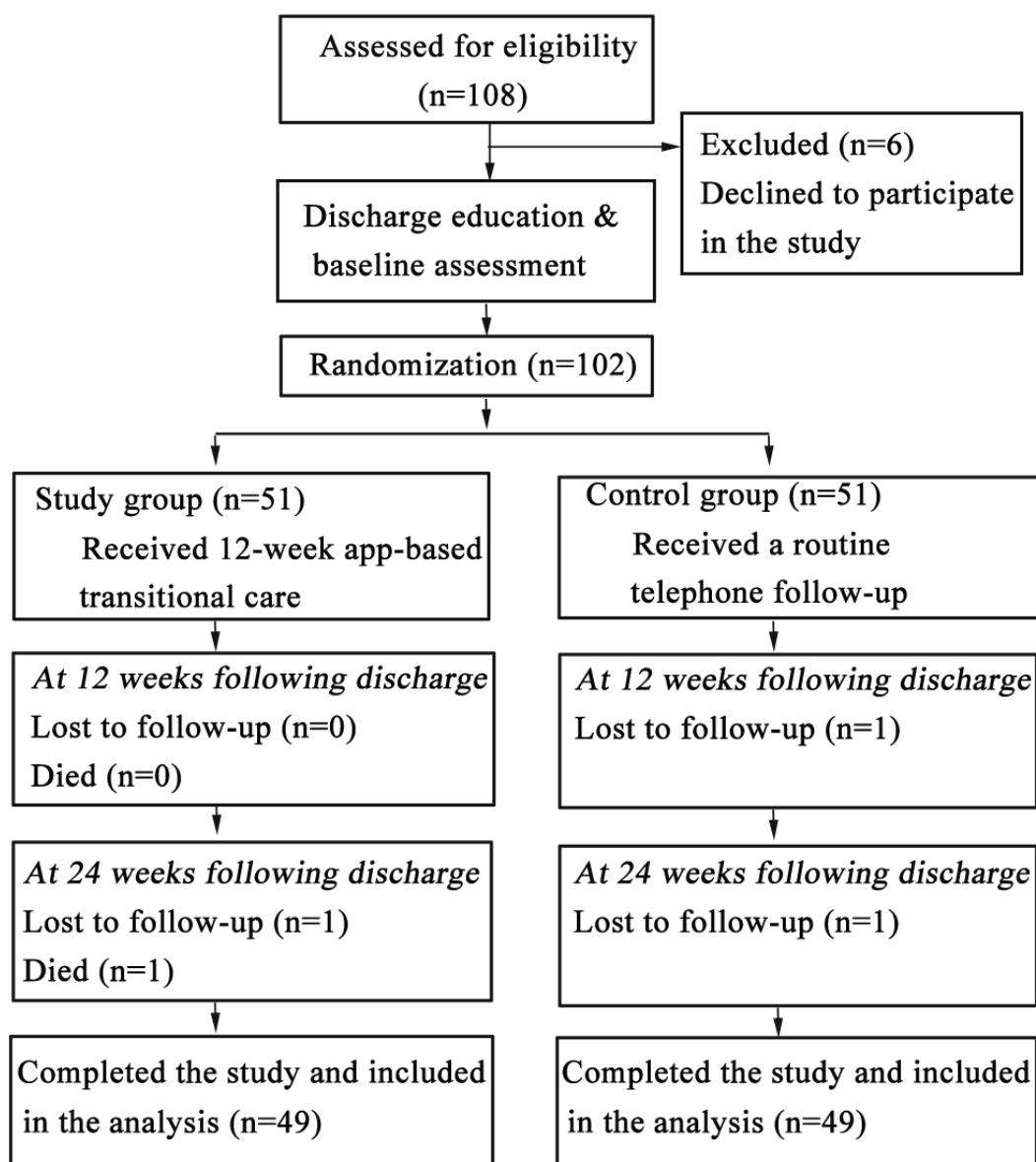
## Statistical Analysis

SPSS version 25.0 software (IBM Corp) was used for data analysis. The frequency, percentage, mean, SD, median, and IQR were used to describe the demographic and disease-related data of the patients in the study group and those in the control group. Mean (SD) or median (IQR) was used to describe the scores of the MSES and SF-36. At the baseline assessment, the chi-squared test,  $t$  test, or Wilcoxon rank-sum test was used for balance testing of all variables, including demographic disease characteristics and outcome measures, between the study and control groups. The normality of the outcome indicators was tested and Box-Cox transformation was used to transform nonnormally distributed data into normally distributed data. Repeated measures analysis of variance was used to analyze the main effects of time, group, and time\*group interaction on the MSES and SF-36 scores. In the presence of an interaction effect, simple effect analysis was further performed. A  $P$  value  $< .05$  was considered to indicate a statistically significant difference.

## Results

### Response to Follow-Up

Figure 4 shows the CONSORT (Consolidated Standards of Reporting Trials) flow diagram of this study. Patients with SCI ( $n=108$ ) were assessed for eligibility, and 102 patients meeting the inclusion criteria were randomly allocated to the study group and the control group. Eventually, only 98 patients completed the study (49 patients per group). In the study group, two patients were lost to follow-up at week 24. The reasons were suicide and disconnection in the two patients. In the control group, two patients were lost to follow-up (owing to disconnection) at weeks 12 and 24. The overall effective follow-up rate in this study was 96.1% (98/102), and the attrition rate was 3.9% (4/102).

**Figure 4.** The CONSORT flow diagram of this study.

### Balance Test

#### *Between the Group Within Guangzhou and the Group Outside Guangzhou*

There were two research centers within Guangzhou and two outside Guangzhou (in Shiyan and Chengdu). Because our study researcher team was in Guangzhou, it was easier to control the study quality of research centers within Guangzhou. Therefore, we divided the four centers into two groups according to the

region (the group within Guangzhou and the group outside Guangzhou). As shown in Table 2, there were no statistically significant differences in the demographic and disease-related data and the baseline scores of self-efficacy and QOL among patients between the group within Guangzhou and the group outside Guangzhou. Thus, we regarded the patients in the four research centers as a whole and carried out repeated measures analysis of variance to analyze the outcome indicators between the study group and control group.

**Table 2.** Balance test between the group within Guangzhou and the group outside Guangzhou (T<sub>0</sub>).

Variable	Group within Guangzhou (n=34)	Group outside Guangzhou (n=64)	Total (n=98)	$t/\chi^2/Z$	P value
Age (years), mean (SD)	41.62 (13.51)	41.77 (11.46)	41.71 (12.14)	$t_{96}=-0.06$	.96 <sup>a</sup>
<b>Sex, n (%)</b>				$\chi^2_1=0.25$	.62 <sup>b</sup>
Male	29 (85)	52 (81)	81 (83)		
Female	5 (15)	12 (19)	17 (17)		
<b>Education level, n (%)</b>				$\chi^2_3=1.40$	.71 <sup>b</sup>
Primary school	9 (27)	19 (30)	28 (29)		
Junior high school	13 (38)	27 (42)	40 (41)		
High school	8 (24)	9 (14)	17 (17)		
College or above	4 (12)	9 (14)	13 (13)		
<b>Marital status, n (%)</b>				$\chi^2_1=0.79$	.37 <sup>b</sup>
Unmarried/divorced	4 (12)	12 (19)	16 (16)		
Married	30 (88)	52 (81)	82 (84)		
<b>Income (RMB/person/month), n (%)</b>				$\chi^2_2=5.63$	.06 <sup>b</sup>
<3000	18 (53)	47 (73)	65 (66)		
3000-4999	9 (27)	13 (20)	22 (22)		
≥5000	7 (21)	4 (6)	11 (11)		
<b>Occupation, n (%)</b>				N/A <sup>c</sup>	.57 <sup>d</sup>
Employed	2 (6)	3 (5)	5 (5)		
Unemployed	32 (94)	61 (95)	93 (95)		
<b>Disease duration (months), n (%)</b>				$\chi^2_1=0.52$	.47 <sup>b</sup>
0-12	29 (85)	59 (92)	88 (90)		
13-22	5 (15)	5 (8)	10 (10)		
<b>Etiology, n (%)</b>				N/A	>.99 <sup>d</sup>
Traumatic	31 (91)	58 (91)	89 (91)		
Nontraumatic	3 (9)	6 (9)	9 (9)		
<b>Injury severity, n (%)</b>				$\chi^2_1=0.08$	.78 <sup>b</sup>
Complete SCI <sup>e</sup>	16 (47)	32 (50)	48 (49)		
Incomplete SCI	18 (53)	32 (50)	50 (51)		
<b>Injury level, n (%)</b>				$\chi^2_2=0.17$	.92 <sup>b</sup>
Cervical	11 (32)	18 (28)	29 (30)		
Thoracic	15 (44)	31 (48)	46 (47)		
Lumbar/sacral/cauda equina	8 (24)	15 (23)	23 (24)		
<b>MSES<sup>f</sup> score, mean (SD)</b>				$t_{96}=-1.47$	.15 <sup>a</sup>
Social function self-efficacy	67.80 (20.68)	74.06 (21.64)	70.93 (21.29)		
General self-efficacy	23.55 (6.73)	25.14 (6.75)	24.35 (6.75)	$t_{96}=-1.17$	.25 <sup>a</sup>
General self-efficacy	16.25 (5.75)	17.84 (6.21)	17.04 (6.00)	$t_{96}=-1.32$	.19 <sup>a</sup>

Variable	Group within Guangzhou (n=34)	Group outside Guangzhou (n=64)	Total (n=98)	$t/\chi^2/Z$	P value
Person function self-efficacy	28.00 (9.79)	31.08 (11.01)	29.54 (10.48)	$t_{96}=-1.46$	.15 <sup>a</sup>
<b>SF-36<sup>g</sup> score, median (IQR)</b>	58.69 (52.29-81.16)	57.69 (49.89-68.48)	58.16 (50.53-71.30)	$Z=-1.23$	.22 <sup>h</sup>
PCS <sup>i</sup>	20.28 (17.95-29.91)	18.06 (13.41-21.63)	19.17 (14.64-26.40)	$Z=-0.90$	.37 <sup>h</sup>
MCS <sup>j</sup>	36.78 (30.95-52.76)	40.65 (30.69-45.63)	38.63 (30.95-46.04)	$Z=-0.01$	.99 <sup>h</sup>

<sup>a</sup>P value of the independent sample *t* test between the study and control groups.

<sup>b</sup>P value of the chi-square test between the study and control groups.

<sup>c</sup>N/A: not applicable.

<sup>d</sup>P value of the Fisher exact test between the study and control groups.

<sup>e</sup>SCI: spinal cord injury.

<sup>f</sup>MSES: Moorong Self-Efficacy Scale.

<sup>g</sup>SF-36: 36-item Short-Form Health Survey.

<sup>h</sup>P value of the Wilcoxon rank-sum test between the study and control groups.

<sup>i</sup>PCS: physical component summary.

<sup>j</sup>MCS: mental component summary.

### ***Between the Study and Control Groups***

There was no significant difference in baseline data between the study and control groups (Table 3). On average, the majority of patients were male (81/98, 83%) with a mean age of 41.71 (SD 12.14) years (range 18-65 years). Most patients had an education level of junior high school or below (68/98, 69%). Almost 95% (93/98) of patients were unemployed. Prior to the injury, the majority were workers and farmers (workers: 55/98, 56%; farmers: 19/98, 19%).

The disease duration in most patients was less than 1 year (88/98, 90%) (range 1-22 months). Traumatic SCI accounted for 91% (89/98) of cases. Among them, the top three causes of the disease were falling from a height, traffic accidents, and being struck by objects (42/98, 43%; 31/98, 32%; and 13/98, 13%; respectively). The numbers of patients with complete and incomplete SCI were almost equal. Thoracic SCI was the most common injury (46/98, 47%).

**Table 3.** Balance test between the study and control groups.

Variable	Study group (n=49)	Control group (n=49)	Total (n=98)	$t/\chi^2/Z$	P value
Age (years), mean (SD)	40.37 (12.18)	43.06 (12.06)	41.71 (12.14)	$t_{96}=-1.10$	.27 <sup>a</sup>
<b>Sex, n (%)</b>				$\chi^2_1=0.07$	.79 <sup>b</sup>
Male	41 (84)	40 (82)	81 (83)		
Female	8 (16)	9 (18)	17 (17)		
<b>Education level, n (%)</b>				$\chi^2_3=4.15$	.25 <sup>b</sup>
Primary school	16 (33)	12 (25)	28 (29)		
Junior high school	20 (41)	20 (41)	40 (41)		
High school	5 (10)	12 (25)	15 (17)		
College or above	8 (16)	5 (10)	10 (11)		
<b>Marital status, n (%)</b>				$\chi^2_1=0.00$	>.99 <sup>b</sup>
Unmarried/divorced	8 (16)	8 (16)	16 (16)		
Married	41 (84)	41 (84)	82 (84)		
<b>Income (RMB/person/month), n (%)</b>				$\chi^2_2=1.39$	.50 <sup>b</sup>
<3000	35 (71)	30 (61)	65 (66)		
3000-4999	10 (20)	12 (25)	22 (22)		
≥5000	4 (8)	7 (14)	11 (11)		
<b>Occupation, n (%)</b>				N/A <sup>c</sup>	>.99 <sup>d</sup>
Employed	3 (6)	2 (4)	5 (5)		
Unemployed	46 (94)	47 (96)	93 (95)		
<b>Disease duration (months), n (%)</b>				$\chi^2_1=0.45$	.51 <sup>b</sup>
0-12	45 (92)	43 (88)	88 (90)		
13-22	4 (8)	6 (12)	10 (10)		
<b>Etiology, n (%)</b>				N/A	>.99 <sup>d</sup>
Traumatic	45 (92)	44 (90)	89 (91)		
Nontraumatic	4 (8)	5 (10)	9 (9)		
<b>Injury severity, n (%)</b>				$\chi^2_1=0.16$	.69 <sup>b</sup>
Complete SCI <sup>e</sup>	25 (51)	23 (47)	48 (49)		
Incomplete SCI	24 (49)	26 (53)	50 (51)		
<b>Injury level, n (%)</b>				$\chi^2_2=0.17$	.92 <sup>b</sup>
Cervical	14 (29)	15 (31)	29 (30)		
Thoracic	24 (49)	22 (45)	46 (47)		
Lumbar/sacral/cauda equina	11 (22)	12 (25)	23 (24)		
<b>MSES<sup>f</sup> score, mean (SD)</b>					
Social function self-efficacy	67.80 (20.68)	74.06 (21.64)	70.93 (21.29)	$t_{96}=-1.47$	.15 <sup>a</sup>
General self-efficacy	23.55 (6.73)	25.14 (6.75)	24.35 (6.75)	$t_{96}=-1.17$	.25 <sup>a</sup>
Person function self-efficacy	16.25 (5.75)	17.84 (6.21)	17.04 (6.00)	$t_{96}=-1.32$	.19 <sup>a</sup>
Person function self-efficacy	28.00 (9.79)	31.08 (11.01)	29.54 (10.48)	$t_{96}=-1.46$	.15 <sup>a</sup>

Variable	Study group (n=49)	Control group (n=49)	Total (n=98)	$t/\chi^2/Z$	<i>P</i> value
<b>SF-36<sup>g</sup> score, median (IQR)</b>	57.55 (49.71-67.07)	59.50 (51.44-71.69)	58.16 (50.53-71.30)	$Z=-0.71$	.48 <sup>h</sup>
PCS <sup>i</sup>	19.38 (14.14-25.11)	18.70 (14.94-27.20)	19.17 (14.64-26.40)	$Z=-0.05$	.96 <sup>h</sup>
MCS <sup>j</sup>	38.27 (31.40-45.02)	38.85 (30.02-48.08)	38.63 (30.95-46.04)	$Z=-0.49$	.62 <sup>h</sup>

<sup>a</sup>*P* value of the independent sample *t* test between the study and control groups.

<sup>b</sup>*P* value of the chi-square test between the study and control groups.

<sup>c</sup>N/A: not applicable.

<sup>d</sup>*P* value of the Fisher exact test between the study and control groups.

<sup>e</sup>SCI: spinal cord injury.

<sup>f</sup>MSES: Moorong Self-Efficacy Scale.

<sup>g</sup>SF-36: 36-item Short-Form Health Survey.

<sup>h</sup>*P* value of the Wilcoxon rank-sum test between the study and control groups.

<sup>i</sup>PCS: physical component summary.

<sup>j</sup>MCS: mental component summary.

### MSES in the Study and Control Groups

The total scores and three-factor structure scores of the MSES for the study and control groups at T<sub>0</sub>, T<sub>1</sub>, and T<sub>2</sub> are shown in Table 4. There were no differences between the groups in time effects and group effects. However, differences in the interaction effects were statistically significant (total scores:  $F_{2,95}=20.389$ ,  $P<.001$ ; social function self-efficacy:  $F_{2,95}=13.445$ ,  $P<.001$ ; general self-efficacy:  $F_{2,95}=16.063$ ,  $P<.001$ ; person function self-efficacy:  $F_{2,95}=13.604$ ,  $P<.001$ ). Simple effect analysis was

applied to further compare the scores of the study and control groups (Table 5). The total scores and three-factor structure scores at T<sub>2</sub> were significantly higher in the study group than in the control group (total scores:  $F_1=8.506$ ,  $P=.004$ ; social function self-efficacy:  $F_1=8.698$ ,  $P=.003$ ; general self-efficacy:  $F_1=6.684$ ,  $P=.01$ ; person function self-efficacy:  $F_1=6.684$ ,  $P=.01$ ). Trend charts (Figure 5) showed that the total score and scores of the three-factor structures in the study group trended upward over time. In contrast, the scores in the control group showed a decreasing trend over time.

**Table 4.** Repeated measures analysis of variance for the Moorong Self-Efficacy Scale and the 36-item Short-Form Health Survey between the study group (n=49) and control group (n=49).

Outcome	T <sub>0</sub> <sup>a</sup>	T <sub>1</sub> <sup>b</sup>	T <sub>2</sub> <sup>c</sup>	Time effect		Group effect		Interaction effect	
				F <sub>2,95</sub>	P value	F <sub>1,96</sub>	P value	F <sub>2,95</sub>	P value
<b>MSES<sup>d</sup> score<sup>e</sup>, mean (SD)</b>				1.556	.22	0.144	.71	20.389	<.001
Study group	67.80 (20.68)	71.90 (20.36)	76.29 (19.50)						
Control group	74.06 (21.64)	73.33 (18.46)	64.49 (19.33)						
<b>Social function self-efficacy<sup>e</sup>, mean (SD)</b>				1.234	.30	0.275	.60	13.445	<.001
Study group	23.55 (6.73)	24.00 (6.49)	25.45 (6.06)						
Control group	25.14 (6.75)	24.45 (5.46)	21.69 (6.21)						
<b>General self-efficacy<sup>e</sup>, mean (SD)</b>				0.607	.55	0.001	.97	16.063	<.001
Study group	16.25 (5.75)	16.78 (5.56)	18.38 (5.46)						
Control group	17.84 (6.21)	17.98 (5.42)	15.47 (5.64)						
<b>Person function self-efficacy<sup>e</sup>, mean (SD)</b>				1.745	.18	0.193	.66	13.604	<.001
Study group	28.00 (9.79)	31.12 (9.62)	32.47 (9.57)						
Control group	31.08 (11.01)	30.90 (9.65)	27.33 (9.36)						
<b>SF-36<sup>f</sup> score<sup>g</sup>, median (IQR)</b>				6.671	.002	0.082	.78	3.059	.05
Study group	57.55 (49.71-67.07)	61.06 (52.23-74.67)	65.36 (59.98-78.54)						
Control group	59.50 (51.44-71.69)	58.83 (53.24-71.74)	58.77 (53.37-79.19)						
<b>PCS<sup>h</sup> score<sup>g</sup>, median (IQR)</b>				24.516	<.001	0.459	.50	2.551	.08
Study group	19.38 (14.14-25.11)	27.20 (19.37-33.50)	23.92 (20.32-32.80)						
Control group	18.70 (14.94-27.20)	25.66 (20.13-31.06)	20.80 (15.11-28.12)						
<b>MCS<sup>i</sup> score<sup>g</sup>, median (IQR)</b>				2.039	.14	0.086	.77	0.421	.66
Study group	38.27 (31.40-45.02)	30.60 (26.17-49.58)	40.01 (32.03-49.59)						
Control group	38.85 (30.02-48.08)	31.79 (27.84-40.42)	41.22 (31.38-53.57)						

<sup>a</sup>Prior to discharge (baseline).<sup>b</sup>At 12 weeks following discharge.<sup>c</sup>At 24 weeks following discharge.<sup>d</sup>MSES: Moorong Self-Efficacy Scale.<sup>e</sup>Normally distributed data.<sup>f</sup>SF-36: 36-item Short-Form Health Survey.<sup>g</sup>Nonnormally distributed data. After conversion to normally distributed data by Box-Cox transform, the data were subsequently analyzed using repeated measures analysis of variance.<sup>h</sup>PCS: physical component summary.<sup>i</sup>MCS: mental component summary.



**Table 5.** Simple effect analysis results of the Moorong Self-Efficacy Scale between the study and control groups.

Outcome	Time	df	Mean square	F	P value
MSES <sup>a</sup>	T <sub>0</sub> <sup>b</sup>	1	961.724	2.400	.12
MSES	T <sub>1</sub> <sup>c</sup>	1	50.000	0.125	.72
MSES	T <sub>2</sub> <sup>d</sup>	1	3409.020	8.506	.004
Social function self-efficacy	T <sub>0</sub>	1	62.082	1.563	.21
Social function self-efficacy	T <sub>1</sub>	1	4.939	0.124	.73
Social function self-efficacy	T <sub>2</sub>	1	345.469	8.698	.003
General self-efficacy	T <sub>0</sub>	1	62.082	1.926	.17
General self-efficacy	T <sub>1</sub>	1	35.520	1.102	.30
General self-efficacy	T <sub>2</sub>	1	205.755	6.382	.01
Person function self-efficacy	T <sub>0</sub>	1	232.663	2.400	.12
Person function self-efficacy	T <sub>1</sub>	1	1.235	0.013	.91
Person function self-efficacy	T <sub>2</sub>	1	648.000	6.684	.01

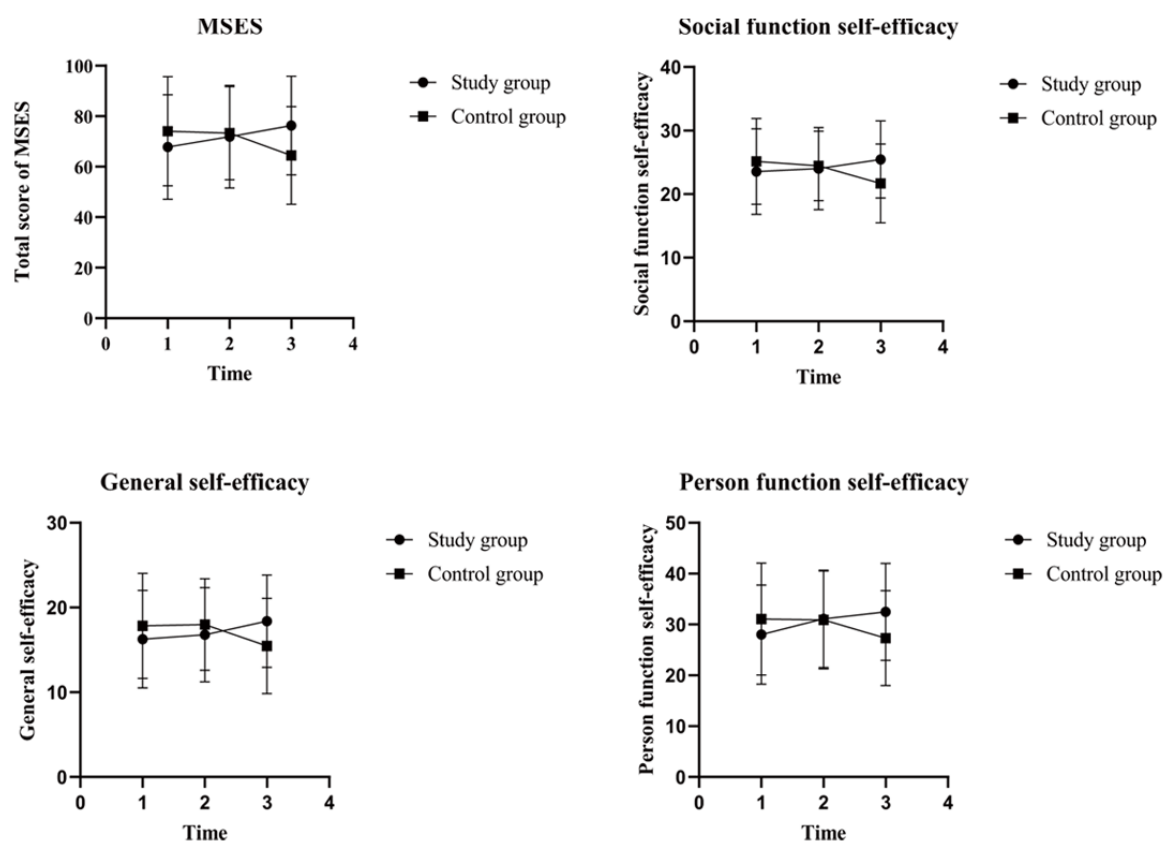
<sup>a</sup>MSES: Moorong Self-Efficacy Scale.

<sup>b</sup>Prior to discharge (baseline).

<sup>c</sup>At 12 weeks following discharge.

<sup>d</sup>At 24 weeks following discharge.

**Figure 5.** Trend charts of total Moorong Self-efficacy Scale (MSES) scores and scores of three-factor structures. Time 1: prior to discharge (baseline); time 2: at 12 weeks following discharge; time 3: at 24 weeks following discharge.

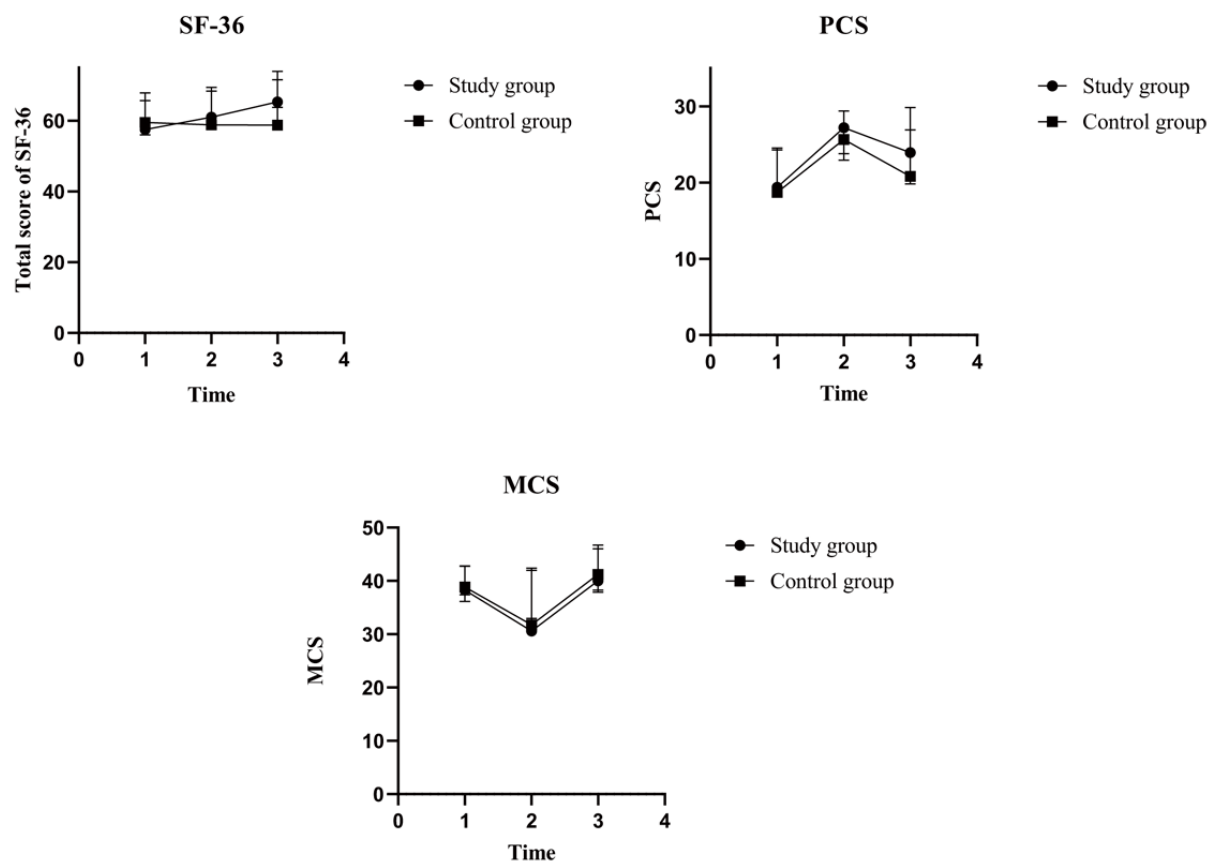


### SF-36 in the Study and Control Groups

The total scores and two-component scores of the SF-36 for the study and control groups at  $T_0$ ,  $T_1$ , and  $T_2$  are summarized in Table 4. Differences in the time effects of the total scores and PCS scores were statistically significant (total scores:  $F_{2,95}=6.671$ ,  $P=.002$ ; PCS:  $F_{2,95}=24.516$ ,  $P<.001$ ). However, there were no differences between the study and control groups

in the time effects in the MCS, interaction effects, and group effects. Trend charts (Figure 6) showed that the total scores in the study group trended upward, whereas the control group scores decreased over time. The PCS scores of both groups showed an increasing tendency at the beginning and then a declining tendency after 12 weeks following discharge. The tendency in MCS scores for both groups was consistent with the opposite tendency in PCS scores.

**Figure 6.** Trend charts of total 36-item Short-Form Health Survey (SF-36) scores and scores of two components. MCS: mental component summary; PCS: physical component summary; time 1: prior to discharge (baseline); time 2: at 12 weeks following discharge; time 3: at 24 weeks following discharge.



## Discussion

### Principal Findings

We found that patients in the study group who received app-based transitional care had better levels of self-efficacy than those in the control group. However, there was no significant statistical difference in QOL between the study and control groups.

Our findings indicated that app-based transitional care was effective at enhancing the self-efficacy of patients with SCI. This is consistent with the results of studies involving other populations, such as puerperium women [18], gestational diabetes patients [49], and stoma patients [50]. The study conducted by MacGillivray et al [19] also revealed that a mobile app could enhance the confidence of patients with SCI in bowel self-management. The study showed that reinforcing the

health-related knowledge of patients can enhance their self-efficacy and self-management behaviors [51]. In our study, the follow-up indicators were extracted from the ICF components, body function, body structure, activities and participation, and contextual factors, which covered the most important issues faced by home-dwelling patients with SCI. These issues included physiological and psychological functional disorders, complications, activities of daily living, and social participation. According to the assessment results of each indicator, standardized health education was provided to patients with SCI, focusing on their knowledge, attitudes, and practice. The delivery and reinforcement of health information from medical staff to patients with SCI may be the main reason for the promotion of self-efficacy in patients. Moreover, app-based transitional care extends social support from professional medical staff to patients. It was demonstrated that higher levels of social support are linked to higher levels of

self-efficacy [52]. Compared with the control group, the social support provided to patients and caregivers in the study group by the multidisciplinary team was more diverse. Additionally, the abilities of caregivers in caring for patients were also assessed, and they received interventions in this study. This approach was also valuable for establishing an effective social support system for the patients.

Although we observed that the study group exhibited a greater improvement in QOL total scores than the control group, there were no significant differences observed between the study and control groups. The results were also consistent with those of the study conducted by Kryger et al [53], which examined the effect of an app on patient QOL. Among other tele-interventions for patients with SCI, Shem et al [54] used FaceTime, a videotelephony app installed on an iPad (Apple Inc), to provide a telemedicine program, and it did not greatly improve the QOL of patients with SCI at 6 months following discharge from the hospital. It is established that the effects of SCI on patients can influence their physiological health, mental health, and social participation in the long term [55]. Therefore, the QOL of patients with SCI was not easily altered. QOL is a multidimensional concept, composed of one's state of the physical, psychological, and social domains [56]. For the physical aspect, the disabilities manifest with paraplegia or tetraplegia, and they seriously affect the functioning and daily activities of patients. For the psychological aspect, patients with SCI often exhibit negative reactions, such as anxiety and depression due to the accident, disabilities, pain, etc, which

greatly affect the subjective feelings of patients regarding life [57]. For social participation, some studies have shown that relevant barriers to social participation exist for patients with SCI owing to environmental constraints, the attitudes of others, and family economic status [58]. Therefore, improving the QOL of patients with SCI requires patients, families, health professionals, and even the whole society to provide long-term integrated support covering all aspects of their physical, mental, and social domains. Notably, the QOL total score and PCS score of the study group showed better trends than those of the control group. Fan [59] conducted an RCT and found that the QOL of patients with SCI at 12 months following discharge improved. A longer observation period may provide more valuable information regarding the effects of app-based transitional care on the QOL of patients with SCI.

### Limitations

The study had some limitations. First, when developing the outcome indicators suitable for SCI follow-up, we did not include patients in our expert panel. In the future, quantitative or qualitative research could be performed to make the follow-up indicators more comprehensive and meet patient needs. Second, the sample size of the study was small. Owing to the low incidence of SCI and the limitation on the disease duration (within 2 years), only 98 eligible patients recruited from four research centers completed the follow-up. Third, the observation period was limited. We followed discharged patients with SCI up to only 24 weeks. Thus, the long-term effects of app-based transitional care should be further investigated.

### Acknowledgments

The authors thank all patients with SCI who participated in this study. In addition, the authors express their sincere gratitude to the software Engineer, Bing Xie, for providing technical support for the development of the app. This study was supported by the National Natural Science Foundation of China (71603293) and the Fundamental Research Funds for the Central Universities (20ykpy88).

### Authors' Contributions

All authors participated in study design and execution, and approved the final version of the submitted paper. TL drafted the manuscript and was responsible for the quality control of the study. SX revised the manuscript and participated in the development of the app and execution of the study. YW, JT, XH, and TY contributed to the execution of the study. KL was responsible for the design, execution of the study, and revision of the manuscript.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.

[PDF File (Adobe PDF File), 1264 KB - [mhealth\\_v9i4e22960\\_app1.pdf](#) ]

### References

1. Cijssouw A, Adriaansen JJE, Tepper M, Dijkstra CA, van Linden S, ALLRISC, et al. Associations between disability-management self-efficacy, participation and life satisfaction in people with long-standing spinal cord injury. *Spinal Cord* 2017 Jan;55(1):47-51. [doi: [10.1038/sc.2016.80](#)] [Medline: [27241451](#)]
2. International Perspectives on Spinal Cord Injury. World Health Organization. 2013. URL: [http://apps.who.int/iris/bitstream/handle/10665/94190/9789241564663\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/94190/9789241564663_eng.pdf?sequence=1) [accessed 2020-06-11]

3. Li J, Liu G, Zheng Y, Hao C, Zhang Y, Wei B, et al. The epidemiological survey of acute traumatic spinal cord injury (ATSCI) of 2002 in Beijing municipality. *Spinal Cord* 2011 Jul;49(7):777-782. [doi: [10.1038/sc.2011.8](https://doi.org/10.1038/sc.2011.8)] [Medline: [21383758](https://pubmed.ncbi.nlm.nih.gov/21383758/)]
4. Pan J, Li X, Zeng C, Qian L, Li LJ, Tan J. Retrospective study of acute spinal cord injury between 2005 and 2007 in Pudong New Area, Shanghai. *Journal of Tongji University of Medical Science* 2009;30(5):131-135 [FREE Full text]
5. Ning G. Study on epidemiology of spinal cord injury in Tianjin. *Tianjin Medical University* 2012;30. [doi: [10.7666/d.Y2232511](https://doi.org/10.7666/d.Y2232511)]
6. Quel de Oliveira C, Refshauge K, Middleton J, de Jong L, Davis GM. Effects of Activity-Based Therapy Interventions on Mobility, Independence, and Quality of Life for People with Spinal Cord Injuries: A Systematic Review and Meta-Analysis. *J Neurotrauma* 2017 May 01;34(9):1726-1743. [doi: [10.1089/neu.2016.4558](https://doi.org/10.1089/neu.2016.4558)] [Medline: [27809702](https://pubmed.ncbi.nlm.nih.gov/27809702/)]
7. Wilson CS, Nassar SL, Ottomanelli L, Barnett SD, Njoh E. Gender differences in depression among veterans with spinal cord injury. *Rehabil Psychol* 2018 May;63(2):221-229. [doi: [10.1037/rep0000221](https://doi.org/10.1037/rep0000221)] [Medline: [29878827](https://pubmed.ncbi.nlm.nih.gov/29878827/)]
8. Yarbrough CK, Gamble PG, Burhan Janjua M, Tang M, Ghenbot R, Zhang AJ, et al. Readmission after spinal cord injury: analysis of an institutional cohort of 795 patients. *J Neurosurg Sci* 2018 Jun;62(3):265-270. [doi: [10.23736/S0390-5616.16.03664-X](https://doi.org/10.23736/S0390-5616.16.03664-X)] [Medline: [27152452](https://pubmed.ncbi.nlm.nih.gov/27152452/)]
9. Sun W. Study on self-care ability of paraplegia after spinal cord injuries. *Wanfang Data*. 2007. URL: <http://d.wanfangdata.com.cn/thesis/Y1180109> [accessed 2021-03-21]
10. Li P, Fu W. Application of SWOT analysis in extended care for the aged discharged patients in community. *Journal of Nursing Science* 2010;25(05):81-83. [doi: [10.3870/hlxzz.2010.05.081](https://doi.org/10.3870/hlxzz.2010.05.081)]
11. Wan Q, Yao L, Hou S. A qualitative study of the current status and existing problems of hospital-based transitional care. *Chinese Nursing Management* 2013;13(10):38-41 [FREE Full text] [doi: [10.3969/j.issn.1672-1756.2013.010.015](https://doi.org/10.3969/j.issn.1672-1756.2013.010.015)]
12. Tian Y. Advances in research on family visit nursing for chronic diseases in the community. *Qingdao Medical Journal* 2018;50(2):145-147. [doi: [10.3969/j.issn.1006-5571.2018.02.024](https://doi.org/10.3969/j.issn.1006-5571.2018.02.024)]
13. Chen Y, Liu F, Zhang G. Effect of three transitional care methods on the rehabilitation and life quality in patients with spinal cord injury. *Chinese Journal of Spine and Spinal Cord* 2016;26(8):741-744. [doi: [10.3969/j.issn.1004-406X.2016.08.12](https://doi.org/10.3969/j.issn.1004-406X.2016.08.12)]
14. Xu Q. The analysis of mobile applications for smoking cessation based on iOS system. *Wanfang Data*. 2014. URL: <https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMTAzMDISCfkyNjkwNjY3GghweTY2eWNkcA%3D%3D>
15. Buis LR, Dawood K, Kadri R, Dawood R, Richardson CR, Djuric Z, et al. Improving Blood Pressure Among African Americans With Hypertension Using a Mobile Health Approach (the MI-BP App): Protocol for a Randomized Controlled Trial. *JMIR Res Protoc* 2019 Jan 25;8(1):e12601 [FREE Full text] [doi: [10.2196/12601](https://doi.org/10.2196/12601)] [Medline: [30681965](https://pubmed.ncbi.nlm.nih.gov/30681965/)]
16. Dahlberg K, Jaensson M, Nilsson U. "Let the patient decide" - Person-centered postoperative follow-up contacts, initiated via a phone app after day surgery: Secondary analysis of a randomized controlled trial. *Int J Surg* 2019 Jan;61:33-37 [FREE Full text] [doi: [10.1016/j.ijssu.2018.11.022](https://doi.org/10.1016/j.ijssu.2018.11.022)] [Medline: [30503602](https://pubmed.ncbi.nlm.nih.gov/30503602/)]
17. De La Cruz Monroy MFI, Mosahebi A. The Use of Smartphone Applications (Apps) for Enhancing Communication With Surgical Patients: A Systematic Review of the Literature. *Surg Innov* 2019 Apr;26(2):244-259. [doi: [10.1177/1553350618819517](https://doi.org/10.1177/1553350618819517)] [Medline: [30602332](https://pubmed.ncbi.nlm.nih.gov/30602332/)]
18. Zhang X, Wei L, Gu R. Application of Mobile Client APP in Continuous Nursing of Puerperium Women. *Journal of Qilu Nursing* 2017;23(12). [doi: [10.3969/j.issn.1006-7256.2017.12.002](https://doi.org/10.3969/j.issn.1006-7256.2017.12.002)]
19. MacGillivray MK, Sadeghi M, Mills PB, Adams J, Sawatzky BJ, Mortenson WB. Implementing a self-management mobile app for spinal cord injury during inpatient rehabilitation and following community discharge: A feasibility study. *J Spinal Cord Med* 2020 Sep;43(5):676-684 [FREE Full text] [doi: [10.1080/10790268.2019.1614343](https://doi.org/10.1080/10790268.2019.1614343)] [Medline: [31091160](https://pubmed.ncbi.nlm.nih.gov/31091160/)]
20. Dolbow DR, Gorgey AS, Ketchum JM, Gater DR. Home-based functional electrical stimulation cycling enhances quality of life in individuals with spinal cord injury. *Top Spinal Cord Inj Rehabil* 2013;19(4):324-329 [FREE Full text] [doi: [10.1310/sci1904-324](https://doi.org/10.1310/sci1904-324)] [Medline: [24244097](https://pubmed.ncbi.nlm.nih.gov/24244097/)]
21. Park C, Otober E, Ullman J, Rogers J, Fasihuddin F, Garg S, et al. Impact on Readmission Reduction Among Heart Failure Patients Using Digital Health Monitoring: Feasibility and Adoptability Study. *JMIR Med Inform* 2019 Nov 15;7(4):e13353 [FREE Full text] [doi: [10.2196/13353](https://doi.org/10.2196/13353)] [Medline: [31730039](https://pubmed.ncbi.nlm.nih.gov/31730039/)]
22. Middleton J, Tran Y, Craig A. Relationship between quality of life and self-efficacy in persons with spinal cord injuries. *Arch Phys Med Rehabil* 2007 Dec;88(12):1643-1648. [doi: [10.1016/j.apmr.2007.09.001](https://doi.org/10.1016/j.apmr.2007.09.001)] [Medline: [18047880](https://pubmed.ncbi.nlm.nih.gov/18047880/)]
23. Baldassin V, Shimizu HE, Fachin-Martins E. Computer assistive technology and associations with quality of life for individuals with spinal cord injury: a systematic review. *Qual Life Res* 2018 Mar;27(3):597-607. [doi: [10.1007/s11136-018-1804-9](https://doi.org/10.1007/s11136-018-1804-9)] [Medline: [29417427](https://pubmed.ncbi.nlm.nih.gov/29417427/)]
24. Gupta A, Deepika S, Taly AB, Srivastava A, Surender V, Thyloth M. Quality of life and psychological problems in patients undergoing neurological rehabilitation. *Ann Indian Acad Neurol* 2008 Oct;11(4):225-230 [FREE Full text] [doi: [10.4103/0972-2327.44557](https://doi.org/10.4103/0972-2327.44557)] [Medline: [19893678](https://pubmed.ncbi.nlm.nih.gov/19893678/)]
25. Bandura A. Self-efficacy: The exercise of control. New York, NY, USA: WH Freeman and Company; 1997.
26. Middleton JW, Tran Y, Lo C, Craig A. Reexamining the Validity and Dimensionality of the Moorong Self-Efficacy Scale: Improving Its Clinical Utility. *Arch Phys Med Rehabil* 2016 Dec;97(12):2130-2136. [doi: [10.1016/j.apmr.2016.05.027](https://doi.org/10.1016/j.apmr.2016.05.027)] [Medline: [27422349](https://pubmed.ncbi.nlm.nih.gov/27422349/)]

27. Craig A, Tran Y, Guest R, Middleton J. Trajectories of Self-Efficacy and Depressed Mood and Their Relationship in the First 12 Months Following Spinal Cord Injury. *Arch Phys Med Rehabil* 2019 Mar;100(3):441-447. [doi: [10.1016/j.apmr.2018.07.442](https://doi.org/10.1016/j.apmr.2018.07.442)] [Medline: [30218640](https://pubmed.ncbi.nlm.nih.gov/30218640/)]
28. Ginis KAM, Latimer AE, Arbour-Nicitopoulos KP, Bassett RL, Wolfe DL, Hanna SE. Determinants of physical activity among people with spinal cord injury: a test of social cognitive theory. *Ann Behav Med* 2011 Aug;42(1):127-133. [doi: [10.1007/s12160-011-9278-9](https://doi.org/10.1007/s12160-011-9278-9)] [Medline: [21544701](https://pubmed.ncbi.nlm.nih.gov/21544701/)]
29. Bisson T, Newsam CJ. Short-duration, high-intensity bouts of physical therapy to increase self-efficacy, confidence, and function in an individual with incomplete spinal cord injury: A case report. *Physiother Theory Pract* 2017 Nov;33(11):888-895. [doi: [10.1080/09593985.2017.1357155](https://doi.org/10.1080/09593985.2017.1357155)] [Medline: [28792298](https://pubmed.ncbi.nlm.nih.gov/28792298/)]
30. Sun Yat-sen University. Distance transitional care system for patients with spinal cord injury. Copyright Protection Center of China Nov2018SR904705. 2018. URL: <http://www.copyright.com/>
31. Ting L, Kun L, Sumei X. Development of the app for transitional care of people with spinal cord injury based on the ICF. *Chinese Journal of Rehabilitation Medicine* 2019;34(2):154-158. [doi: [10.3969/j.issn.1001-1242.2019.02.006](https://doi.org/10.3969/j.issn.1001-1242.2019.02.006)]
32. Li K, Xie S, Wang Y, Tang J, He X, Liu T, et al. Outcome indicators in the transitional care of people with spinal cord injury in China: a Delphi survey based on the International Classification of Functioning, Disability and Health. *Disabil Rehabil* 2020 Jun 06;42(11):1539-1547. [doi: [10.1080/09638288.2018.1528638](https://doi.org/10.1080/09638288.2018.1528638)] [Medline: [30724613](https://pubmed.ncbi.nlm.nih.gov/30724613/)]
33. Bilbao A, Kennedy C, Chatterji S, Ustün B, Barquero JL, Barth JT. The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *NeuroRehabilitation* 2003;18(3):239-250. [Medline: [14530589](https://pubmed.ncbi.nlm.nih.gov/14530589/)]
34. Wright BD, Mok M. Rasch models overview. *J Appl Meas* 2000;1(1):83-106. [Medline: [12023559](https://pubmed.ncbi.nlm.nih.gov/12023559/)]
35. Jia M, Tang J, Xie S, He X, Wang Y, Liu T, et al. Using a Mobile App-Based International Classification of Functioning, Disability, and Health Set to Assess the Functioning of Spinal Cord Injury Patients: Rasch Analysis. *JMIR Mhealth Uhealth* 2020 Nov 11;8(11):e20723 [FREE Full text] [doi: [10.2196/20723](https://doi.org/10.2196/20723)] [Medline: [33174860](https://pubmed.ncbi.nlm.nih.gov/33174860/)]
36. Li K, Yan T, You L, Xie S, Li Y, Tang J, et al. The inter-rater reliability of the International Classification of Functioning, Disability and Health set for spinal cord injury nursing. *Int J Rehabil Res* 2016 Sep;39(3):240-248. [doi: [10.1097/MRR.000000000000176](https://doi.org/10.1097/MRR.000000000000176)] [Medline: [27171607](https://pubmed.ncbi.nlm.nih.gov/27171607/)]
37. Badran IG. Knowledge, attitude and practice the three pillars of excellence and wisdom: a place in the medical profession. *East Mediterr Health J* 1995;1(1):8-16. [doi: [10.26719/1995.1.1.8](https://doi.org/10.26719/1995.1.1.8)]
38. International standards for neurological classification of spinal cord injury (ISNCSCI). America Spinal Cord Association. 2016. URL: [https://asia-spinalinjury.org/wp-content/uploads/2019/10/ASIA-ISCOS-Worksheet\\_10.2019\\_PRINT-Page-1-2.pdf](https://asia-spinalinjury.org/wp-content/uploads/2019/10/ASIA-ISCOS-Worksheet_10.2019_PRINT-Page-1-2.pdf) [accessed 2020-06-11]
39. Research Randomizer (Version 4.0). Randomizer. URL: <http://www.randomizer.org/> [accessed 2020-06-11]
40. Martin Ginis KA, Shaw RB, Stork MJ, Battalova A, McBride CB. Pilot study of a training program to enhance transformational leadership in Spinal Cord Injury Peer Mentors. *Spinal Cord Ser Cases* 2018;4:34 [FREE Full text] [doi: [10.1038/s41394-018-0065-8](https://doi.org/10.1038/s41394-018-0065-8)] [Medline: [29707237](https://pubmed.ncbi.nlm.nih.gov/29707237/)]
41. Mortenson WB, Noreau L, Miller WC. The relationship between and predictors of quality of life after spinal cord injury at 3 and 15 months after discharge. *Spinal Cord* 2010 Jan;48(1):73-79 [FREE Full text] [doi: [10.1038/sc.2009.92](https://doi.org/10.1038/sc.2009.92)] [Medline: [19636329](https://pubmed.ncbi.nlm.nih.gov/19636329/)]
42. Dorstyn D, Mathias J, Denson L, Robertson M. Effectiveness of telephone counseling in managing psychological outcomes after spinal cord injury: a preliminary study. *Arch Phys Med Rehabil* 2012 Nov;93(11):2100-2108. [doi: [10.1016/j.apmr.2012.06.002](https://doi.org/10.1016/j.apmr.2012.06.002)] [Medline: [22705237](https://pubmed.ncbi.nlm.nih.gov/22705237/)]
43. Middleton JW, Tate RL, Geraghty TJ. Self-Efficacy and Spinal Cord Injury: Psychometric Properties of a New Scale. *Rehabilitation Psychology* 2003 Nov;48(4):281-288. [doi: [10.1037/0090-5550.48.4.281](https://doi.org/10.1037/0090-5550.48.4.281)]
44. Miller SM. The measurement of self-efficacy in persons with spinal cord injury: psychometric validation of the moorong self-efficacy scale. *Disabil Rehabil* 2009;31(12):988-993. [doi: [10.1080/09638280802378025](https://doi.org/10.1080/09638280802378025)] [Medline: [19241199](https://pubmed.ncbi.nlm.nih.gov/19241199/)]
45. Rajati F, Ghanbari M, Hasandokht T, Hosseini SY, Akbarzadeh R, Ashtarian H. Persian version of the Moorong Self-Efficacy Scale: psychometric study among subjects with physical disability. *Disabil Rehabil* 2017 Nov;39(23):2436-2445. [doi: [10.1080/09638288.2016.1226404](https://doi.org/10.1080/09638288.2016.1226404)] [Medline: [27767355](https://pubmed.ncbi.nlm.nih.gov/27767355/)]
46. Wilson JR, Hashimoto RE, Dettori JR, Fehlings MG. Spinal cord injury and quality of life: a systematic review of outcome measures. *Evid Based Spine Care J* 2011 Feb;2(1):37-44 [FREE Full text] [doi: [10.1055/s-0030-1267085](https://doi.org/10.1055/s-0030-1267085)] [Medline: [22956935](https://pubmed.ncbi.nlm.nih.gov/22956935/)]
47. Wan C, Fang J, Shi M. A comparative study on measuring quality of life for drug addicts by SF-36 scale. *Chinese Journal of Behavioral Medicine and Brain Science* 1998(4):259-261. [doi: [10.3760/cma.j.issn.1674-6554.1998.04.012](https://doi.org/10.3760/cma.j.issn.1674-6554.1998.04.012)]
48. Forchheimer M, McAweeney M, Tate DG. Use of the SF-36 among persons with spinal cord injury. *Am J Phys Med Rehabil* 2004 May;83(5):390-395. [doi: [10.1097/01.phm.0000124441.78275.c9](https://doi.org/10.1097/01.phm.0000124441.78275.c9)] [Medline: [15100631](https://pubmed.ncbi.nlm.nih.gov/15100631/)]
49. Han T. Dietary Intervention for Gestational Diabetes Patients Based on Smartphone Application. CNKI. 2019. URL: <http://gb.oversea.cnki.net/KCMS/detail/detail.aspx?filename=1019612422.nh&dbcode=CMFD&dbname=CMFD2019> [accessed 2021-03-21]

50. Wang QQ, Zhao J, Huo XR, Wu L, Yang LF, Li JY, et al. Effects of a home care mobile app on the outcomes of discharged patients with a stoma: A randomised controlled trial. *J Clin Nurs* 2018 Oct;27(19-20):3592-3602. [doi: [10.1111/jocn.14515](https://doi.org/10.1111/jocn.14515)] [Medline: [29775491](https://pubmed.ncbi.nlm.nih.gov/29775491/)]
51. Zhang J, Wong FKY, You L, Zheng M, Li Q, Zhang B, et al. Effects of enterostomal nurse telephone follow-up on postoperative adjustment of discharged colostomy patients. *Cancer Nurs* 2013;36(6):419-428. [doi: [10.1097/NCC.0b013e31826fc8eb](https://doi.org/10.1097/NCC.0b013e31826fc8eb)] [Medline: [23051876](https://pubmed.ncbi.nlm.nih.gov/23051876/)]
52. Atobatele KO, Olaleye OA, Fatoye FA, Hamzat TK. Relationships Between Community Reintegration and Clinical and Psychosocial Attributes in Individuals With Spinal Cord Injury in a Nigerian City. *Top Spinal Cord Inj Rehabil* 2018;24(4):306-314 [FREE Full text] [doi: [10.1310/sci16-00055](https://doi.org/10.1310/sci16-00055)] [Medline: [30459493](https://pubmed.ncbi.nlm.nih.gov/30459493/)]
53. Kryger MA, Crytzer TM, Fairman A, Quinby EJ, Karavolis M, Pramana G, et al. The Effect of the Interactive Mobile Health and Rehabilitation System on Health and Psychosocial Outcomes in Spinal Cord Injury: Randomized Controlled Trial. *J Med Internet Res* 2019 Aug 28;21(8):e14305 [FREE Full text] [doi: [10.2196/14305](https://doi.org/10.2196/14305)] [Medline: [31464189](https://pubmed.ncbi.nlm.nih.gov/31464189/)]
54. Shem K, Sechrist SJ, Loomis E, Isaac L. SCiPad: Effective Implementation of Telemedicine Using iPads with Individuals with Spinal Cord Injuries, a Case Series. *Front Med (Lausanne)* 2017;4:58 [FREE Full text] [doi: [10.3389/fmed.2017.00058](https://doi.org/10.3389/fmed.2017.00058)] [Medline: [28611986](https://pubmed.ncbi.nlm.nih.gov/28611986/)]
55. Foongchomcheay A, Eitvipart AC, Kespichayawattana J, Muangngoen M. Quality of life after spinal cord injury in Thai individuals: A mixed-methods study. *Hong Kong Physiother J* 2019 Jun;39(1):35-55 [FREE Full text] [doi: [10.1142/S1013702519500045](https://doi.org/10.1142/S1013702519500045)] [Medline: [31156316](https://pubmed.ncbi.nlm.nih.gov/31156316/)]
56. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. *Social Science & Medicine* 1995 Nov;41(10):1403-1409. [doi: [10.1016/0277-9536\(95\)00112-K](https://doi.org/10.1016/0277-9536(95)00112-K)]
57. Lim S, Shiue Y, Ho C, Yu S, Kao P, Wang J, et al. Anxiety and Depression in Patients with Traumatic Spinal Cord Injury: A Nationwide Population-Based Cohort Study. *PLoS One* 2017;12(1):e0169623 [FREE Full text] [doi: [10.1371/journal.pone.0169623](https://doi.org/10.1371/journal.pone.0169623)] [Medline: [28081205](https://pubmed.ncbi.nlm.nih.gov/28081205/)]
58. Shi H. Research on emotional status and activity and participation for inpatients with spinal cord injury. *Chinese Journal of Rehabilitation Theory and Practice* 2017;23(8):946-949. [doi: [10.3969/j.issn.1006-9771.2017.08.016](https://doi.org/10.3969/j.issn.1006-9771.2017.08.016)]
59. Fan X. The effect of the WeChat joint nursing outpatient follow-up on the quality of life of patients with spinal cord injury. *International Journal of Nursing* 2017;36(24):3430-3333. [doi: [10.3760/cma.j.issn.1673-4351.2017.24.005](https://doi.org/10.3760/cma.j.issn.1673-4351.2017.24.005)]

## Abbreviations

**AD:** autonomic dysreflexia

**CVI:** content validity index

**ICF:** International Classification of Functioning, Disability, and Health

**MCS:** mental component summary

**MSES:** Moorong Self-Efficacy Scale

**PCS:** physical component summary

**QOL:** quality of life

**RCT:** randomized controlled trial

**SCI:** spinal cord injury

**SF-36:** 36-item Short-Form Health Survey

*Edited by L Buis; submitted 03.08.20; peer-reviewed by U Nilsson, G Pramana; comments to author 18.09.20; revised version received 10.11.20; accepted 08.03.21; published 01.04.21.*

*Please cite as:*

*Liu T, Xie S, Wang Y, Tang J, He X, Yan T, Li K*

*Effects of App-Based Transitional Care on the Self-Efficacy and Quality of Life of Patients With Spinal Cord Injury in China: Randomized Controlled Trial*

*JMIR Mhealth Uhealth* 2021;9(4):e22960

URL: <https://mhealth.jmir.org/2021/4/e22960>

doi: [10.2196/22960](https://doi.org/10.2196/22960)

PMID: [33792555](https://pubmed.ncbi.nlm.nih.gov/33792555/)

©Ting Liu, Sumei Xie, Yingmin Wang, Jie Tang, Xiaokuo He, Tiebin Yan, Kun Li. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 01.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Remote Heart Rhythm Monitoring by Photoplethysmography-Based Smartphone Technology After Cardiac Surgery: Prospective Observational Study

Marie Lamberigts<sup>1</sup>, MSc; Lucas Van Hoof<sup>1</sup>, MD; Tine Proesmans<sup>2</sup>, MSc; Pieter Vandervoort<sup>3</sup>, MD, PhD; Lars Grieten<sup>2</sup>, PhD; Peter Haemers<sup>4</sup>, MD, PhD; Filip Rega<sup>1</sup>, MD, PhD

<sup>1</sup>Department of Cardiac Surgery, University Hospitals Leuven, Leuven, Belgium

<sup>2</sup>Qompium NV, Hasselt, Belgium

<sup>3</sup>Department of Cardiology, Ziekenhuis Oost-Limburg, Genk, Belgium

<sup>4</sup>Department of Cardiology, University Hospitals Leuven, Leuven, Belgium

**Corresponding Author:**

Marie Lamberigts, MSc

Department of Cardiac Surgery

University Hospitals Leuven

Herestraat 49

Leuven, 3000

Belgium

Phone: 32 477 09 49 32

Email: [marie.lamberigts@kuleuven.be](mailto:marie.lamberigts@kuleuven.be)

## Abstract

**Background:** Atrial fibrillation (AF) is the most common arrhythmia after cardiac surgery, yet the precise incidence and significance of arrhythmias after discharge home need to be better defined. Photoplethysmography (PPG)-based smartphone apps are promising tools to enable early detection and follow-up of arrhythmias.

**Objective:** By using a PPG-based smartphone app, we aimed to gain more insight into the prevalence of AF and other rhythm-related complications upon discharge home after cardiac surgery and evaluate the implementation of this app into routine clinical care.

**Methods:** In this prospective, single-center trial, patients recovering from cardiac surgery were asked to register their heart rhythm 3 times daily using a Food and Drug Administration–approved PPG-based app, for either 30 or 60 days after discharge home. Patients with permanent AF or a permanent pacemaker were excluded.

**Results:** We included 24 patients (mean age 60.2 years, SD 12 years; 15/23, 65% male) who underwent coronary artery bypass grafting and/or valve surgery. During hospitalization, 39% (9/23) experienced postoperative AF. After discharge, the PPG app reported AF or atrial flutter in 5 patients. While the app notified flutter in 1 patient, this was a false positive, as electrocardiogram revealed a 2nd-degree, 2:1 atrioventricular block necessitating a permanent pacemaker. AF was confirmed in 4 patients (4/23, 17%) and interestingly, was associated with an underlying postoperative complication in 2 participants (pneumonia n=1, pericardial tamponade n=1). A significant increase in the proportion of measurements indicating sinus rhythm was observed when comparing the first to the second month of follow-up ( $P<.001$ ). In the second month of follow-up, compliance was significantly lower with 2.2 (SD 0.7) measurements per day versus 3.0 (SD 0.8) measurements per day in the first month ( $P=.002$ ). The majority of participants (17/23, 74%), as well as the surveyed primary care physicians, experienced positive value by using the app as they felt more involved in the postoperative rehabilitation.

**Conclusions:** Implementation of smartphone-based PPG technology enables detection of AF and other rhythm-related complications after cardiac surgery. An association between AF detection and an underlying complication was found in 2 patients. Therefore, smartphone-based PPG technology may supplement rehabilitation after cardiac surgery by acting as a sentinel for underlying complications, rhythm-related or otherwise.

(*JMIR Mhealth Uhealth* 2021;9(4):e26519) doi:[10.2196/26519](https://doi.org/10.2196/26519)



## KEYWORDS

cardiac surgery; postoperative follow-up; cardiac rehabilitation; postoperative arrhythmias; atrial fibrillation; photoplethysmography; home-monitoring

## Introduction

Postoperative atrial fibrillation (POAF) is one of the most important complications after cardiac surgery, occurring in 10%-65% of patients, depending on their risk profile and surgery type [1-3]. The burden of this complication on health care and the economy is significant, as it leads to elevated stroke risk, higher postoperative mortality, longer hospitalization, and an increase in medical costs [3-5]. AF and other arrhythmias typically occur early after surgery, with the incidence of AF peaking on the second to third postoperative day, correlating with the inflammatory peak [2,6,7]. In studies with extensive follow-up using noninvasive telemetry or invasive event recorders, late-onset POAF occurs in approximately 9% of patients [8-13]. Recurrent POAF is more common, with an estimated recurrence rate of 30% in the first months postoperatively [14,15]. The precise incidence of late-onset or recurrent AF remains uncertain due to discrepancies between previous studies.

The basic principle of photoplethysmography (PPG) technology is the detection by an optical sensor of variations in light intensity, as reflected by or transmitted through tissue. The tissue perfusion with every heartbeat can be registered, and by extrapolating the RR-interval, the underlying heart rhythm can be determined. FibrCheck (Qompium NV, Hasselt, Belgium) is an example of a Conformité Européenne- and Food and Drug administration-approved smartphone app that employs the phone's flashlight and camera, thereby offering a low-budget and widely accessible platform for early diagnosis and close follow-up of AF [16,17]. Previous studies have extensively confirmed its accuracy in detecting AF and the feasibility of use in a primary care setting or as a large-scale screening tool [17-19]. While there are currently no reports on the systematic use of a PPG-based smartphone app after cardiac surgery, such an approach is promising to detect late-onset or recurrent POAF that may otherwise be missed in routine clinical follow-up. On the other hand, the importance and therapeutic implications of these sometimes short and asymptomatic episodes of AF are uncertain. Therefore, the potential added value of PPG technology in postoperative rehabilitation must be evaluated.

The objective of this study was to gain insight into the prevalence of arrhythmias upon discharge home after cardiac surgery by using a PPG-based app. Furthermore, we aimed to evaluate the added value and obstacles to implementation of this app into routine clinical care.

## Methods

### Study Design

This study was constructed as an observational, monocentric cohort study. An a priori sample size calculation was not

performed for this proof-of-concept trial. This study was approved by the Ethics Committee Research of University Hospitals (UZ)/Katholieke Universiteit Leuven (S63159) and was conducted in accordance with the Declaration of Helsinki.

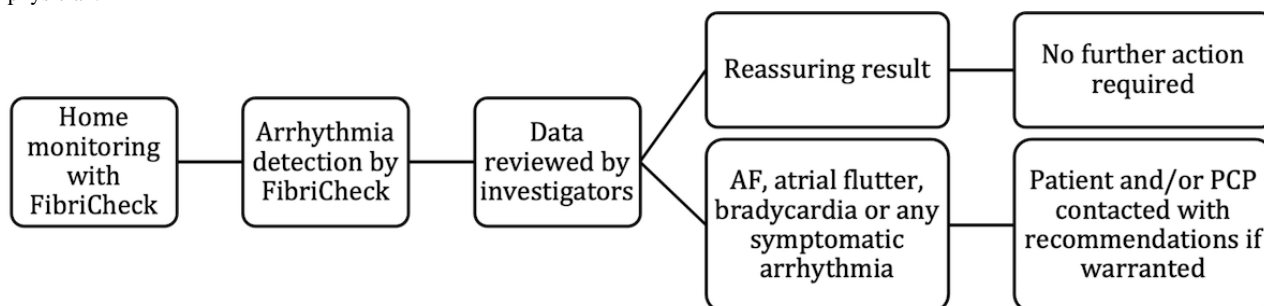
At discharge, participants were educated on the use of the app by a member of the research team before starting a follow-up period of 60 days in which they used FibrCheck to assess their heart rhythm. The home monitoring consisted of 3 measurements per day, with each measurement being 1 minute long. If compliance decreased, the participant was sent a reminder. At the end of the study period, participants were contacted and asked to complete a short questionnaire on their personal experience of using the smartphone app as a follow-up tool.

### FibrCheck App and Evaluation of PPG Registrations

The specific PPG-based smartphone app under evaluation was FibrCheck, which is capable of distinguishing multiple different arrhythmias. Measurements are performed by placing a finger over the camera of a smartphone for 1 minute. During this period, the smartphone's optical sensor detects variations in tissue perfusion related to the heart rhythm, enabling extrapolation of the RR-interval. An example of a measurement is provided in [Multimedia Appendix 1](#). Each measurement is then uploaded to a server and immediately analyzed by a dedicated algorithm, classifying the measurement into 1 of 4 categories (normal, warning, urgent, insufficient quality). Warning measurements include non-AF arrhythmias, and urgent measurements include possible AF arrhythmias. Measurements indicating normal sinus rhythm are validated immediately by the algorithm, whereas measurements with possible irregularities are reviewed by highly trained personnel within, at most, 48 hours. As specified in the instructions for use, every arrhythmia detected by FibrCheck should be confirmed by an electrocardiogram (ECG).

As per standard protocol for FibrCheck, all PPG measurements were immediately available to the research team through the online dashboard of the app. To prevent delays in diagnosis, the investigators checked this dashboard daily for abnormal measurements. If an unconfirmed measurement was suspected by the research team to indicate AF, the study participant was contacted, and the research team suggested that they visit their primary care physician (PCP). Furthermore, patients were instructed at study inclusion to seek emergent medical care if they felt unwell and not rely on the app to guide them. An overview of the protocol in case of arrhythmia detection is shown in [Figure 1](#).

**Figure 1.** Schematic overview of the protocol for management of detected arrhythmias during the follow-up phase. AF: atrial fibrillation; PCP: primary care physician.



## Outcome Measures

The primary outcome was the detection of postoperative arrhythmias after hospital discharge, mainly focusing on late-onset or recurrent postoperative AF. Secondary outcomes included the detection of other rhythm-related events, the impact and added value of FibriCheck on routine postoperative care, and the experience of patients and PCPs.

## Participant Recruitment and Eligibility Criteria

Patients were recruited during their recovery period after cardiac surgery at UZ Leuven. Inclusion criteria consisted of (1) a minimum age of 18 years, (2) possession of a smartphone, (3) underwent cardiac surgery at UZ Leuven, and (4) written informed consent. Exclusion criteria were (1) permanent AF; (2) pacemaker-dependent heart rhythm or ventricular assist device; (3) significant cognitive impairment (eg, dementia); (4) poor finger perfusion (eg, intensive callus formation, pernio); (5) tremor, Parkinson's disease, or other disabilities resulting in the inability to perform measurements; (6) nonnative Dutch speakers; and (7) long and complicated postoperative hospitalization, for example after endocarditis surgery.

## Data Collection and Analysis

Data were collected and processed anonymously, in accordance with regulations on data protection, ethics, and written informed consent. Patient-related information was retrieved from the electronic medical records system used at UZ Leuven, while data on the PPG registrations were exported from the FibriCheck app interface.

Continuous variables are presented as mean (SD), and categorical variables are presented as numbers and percentages. The Shapiro-Wilk test was used to test normality. The Student *t* test, Wilcoxon rank sum test, and Fisher exact test were used to determine significance when comparing variables between subgroups. The 2-proportion *z* test was used to determine the significance of proportions between subgroups. Statistical significance was set at  $P < .05$ , and data were analyzed using R Studio version 1.1.447 (RStudio Inc, Boston, MA) or Microsoft Excel for Mac version 16.36 (Microsoft Corporation, Redmond, WA).

## Results

### Study Population

Between January 29, 2020 and March 12, 2020, patients recovering from cardiac surgery at UZ Leuven were screened for participation. In this period, 24 patients were initially enrolled in the study. A total of 23 participants completed the study period, as 1 participant received a permanent pacemaker for late-onset, second-degree atrioventricular (AV) block with 2:1 conduction. All remaining 23 participants were home monitored for 30 days, and 17 were followed for an additional 30 days. The mean age was 60.2 (SD 12) years, and the majority of the study population was male (15/23, 65%). [Table 1](#) provides an overview of demographic and other variables of the study population.

**Table 1.** Overview of (demographic) variables of the study population (n=23).

Variables	Values
Age (years), mean (SD)	60.2 (12)
Gender (male), n (%)	15 (65)
BMI (kg/m <sup>2</sup> ), mean (SD)	26.5 (5)
Diabetes, n (%)	2 (9)
Hypertension, n (%)	9 (39)
Hypercholesterolemia, n (%)	7 (30)
BMI >30 kg/m <sup>2</sup> , n (%)	3 (13)
CHA <sub>2</sub> DS <sub>2</sub> -VASc <sup>a</sup> score, mean (SD)	2 (1)
HAS-BLED <sup>b</sup> score, mean (SD)	1 (1)
EuroScore II, mean (SD)	1.4 (1)
LVEF <sup>c</sup> , mean (SD)	58.2 (6)
LVDD <sup>d</sup> (mm), mean (SD)	47.9 (8)
LAVI <sup>e</sup> (mL/m <sup>2</sup> ), mean (SD)	30.9 (8)
Length of ICU <sup>f</sup> stay (days) <sup>g</sup> , mean (SD)	1.5 (2)
Total length of stay (days) <sup>g</sup> , mean (SD)	6.9 (3)
<b>Surgery type, n (%)</b>	
AVR <sup>h</sup> /root/arch replacement	8 (35)
CABG <sup>i</sup>	7 (30)
Mitral valve surgery	5 (22)
ASD <sup>j</sup> closure	1 (4)
CABG + AVR	1 (4)
MVP <sup>k</sup> + root replacement	1 (4)

<sup>a</sup>CHA<sub>2</sub>DS<sub>2</sub>-VASc: congestive heart failure, hypertension, age (>75 years), diabetes, stroke/transient ischemic attack, vascular disease, previous myocardial infarction, age (65-75 years), and sex category.

<sup>b</sup>HAS-BLED: hypertension, abnormal renal and liver function, stroke, bleeding, liable INR, elderly, drugs or alcohol.

<sup>c</sup>LVEF: left ventricle ejection fraction.

<sup>d</sup>LVDD: left ventricle diastolic diameter.

<sup>e</sup>LAVI: left atrial volume index.

<sup>f</sup>ICU: intensive care unit.

<sup>g</sup>Uz025 length of ICU stay is removed as an outlier; cut off boundaries were set at >20 days.

<sup>h</sup>AVR: aortic valve repair.

<sup>i</sup>CABG: coronary artery bypass grafting.

<sup>j</sup>ASD: atrium septum defect.

<sup>k</sup>MVP: mitral valve plasty.

During hospitalization, POAF occurred in 9 of the 23 (39%) patients, 4 of whom underwent isolated coronary artery bypass grafting surgery. The remaining 5 participants who experienced in-hospital POAF underwent various procedures. Other arrhythmias detected by telemetry during the postoperative hospitalization included several episodes of supraventricular tachycardia, sinus bradycardia, and frequent ventricular extrasystoles and one ventricular tachycardia episode.

A total of 42 patients were excluded prior to participation, and these patients were significantly older than the study population (average age of 71.7 vs 60.2 years,  $P=.002$ ). Reasons for exclusion were lack of smartphone possession (n=13), patients that were transferred to other hospitals (n=9), permanent pacemaker (n=7), refusal (n=6), permanent AF (n=4), and nonnative Dutch-speaking patients (n=3).

## Overview of All FibrCheck Measurements

The total number of measurements performed by the 23 participants was 3271, with 2808 (2808/3271, 85.9%) being normal, 288 (288/3271, 8.8%) measurements labelled as a warning, and 27 (27/3271, 0.8%) prompted urgent action. The remaining 148 (148/3271, 4.5%) measurements were of insufficient quality. A majority of the warning measurements (159/288, 55.2%) occurred in one patient experiencing very frequent extrasystoles, and 67% (18/27) of all urgent measurements originated from one other patient. Concerning the diagnosis assigned to each measurement, 2792 (2792/3271, 85.4%) were sinus rhythm, 295 (295/3271, 9.0%) measurements were premature ventricular contractions, and bradyarrhythmias and tachyarrhythmias accounted for 5 (5/3271, 0.2%) and 2 (2/3271, 0.1%) of the measurements, respectively. Of all measurements, 27 (27/3271, 0.8%) were diagnosed as atrial flutter or fibrillation.

Most measurements (2929/3271, 89.5%) with the “normal” or “insufficient quality” label were automatically diagnosed as sinus rhythm or “insufficient quality” by the algorithm without further review by FibrCheck staff. The remaining 342 (342/3271, 10.5%) measurements, mostly warning and urgent measurements, were manually reviewed by FibrCheck staff. The mean timespan between measurement performance and review was 16.6 (SD 11.8) hours, and the maximum duration seen was 47.0 hours.

## Detection of Atrial Fibrillation

AF was confirmed in 4 participants (4/23, 17%) and interestingly, was associated with underlying complications in 2 participants. One 70-year-old male participant who

experienced recurrent POAF 7 days after discharge simultaneously presented with a late-occurring pericardial tamponade. After surgical drainage, AF was not reported any more by the app. The second patient with an underlying complication was a 48-year-old man in whom FibrCheck detected late-onset AF on days 4 and 5 after discharge. After subsequent readmission, a diagnosis of pneumonia was made. He received intravenous antibiotics and underwent an electrical cardioversion.

## Detection of a Second-Degree AV Block

During follow-up, FibrCheck indicated bradycardia accompanied by frequent premature ventricular contractions in one participant during the first 9 days after discharge. This participant was a 56-year-old woman with Barlow’s disease who underwent minimally invasive mitral valve repair. Starting on the 10th day after discharge (postoperative day 15), the app indicated atrial flutter. This was found to be a false positive as an ECG revealed a 2nd-degree 2:1 AV block requiring a permanent pacemaker, excluding her for the remainder of the trial.

## First Versus Second Month of Follow-Up

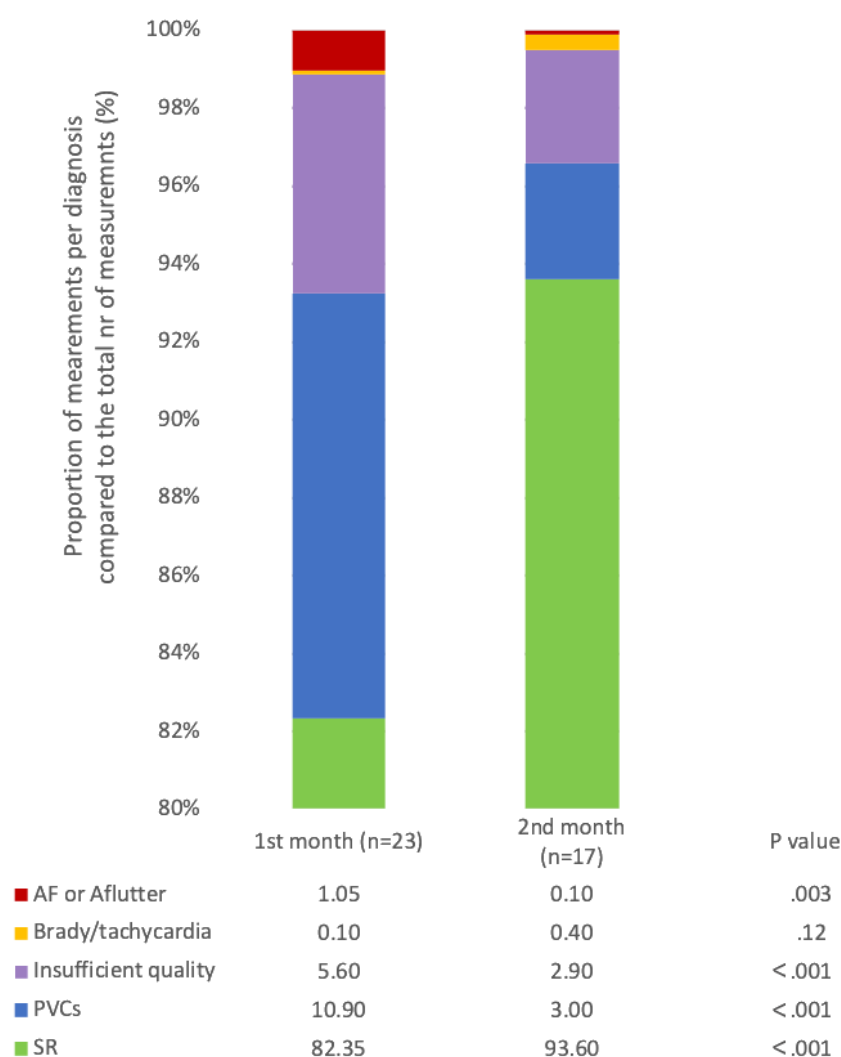
All 23 participants were followed for 30 days, and 17 participants were followed for an additional 30 days. Table 2 depicts the differences between the first and second months of follow-up. Significantly fewer measurements were performed in total, and less were of insufficient quality. We observed a significant increase in the proportion of measurements diagnosed as sinus rhythm as well as a decrease in the prevalence of other arrhythmias (Figure 2).

**Table 2.** Overview of compliance, calculated as the average amount of measurements performed per day, with optimal compliance considered as 3 measurements daily.

Variable	Entire study (n=23)	First month (n=23)	Second month (n=17)	P value <sup>a</sup>
Measurements per day, mean (SD)	2.7 (0.9)	3.0 (0.8)	2.2 (0.7)	.002
Participants performing $\geq 3$ measurements per day, n (%)	7 (30)	12 (52)	0 (0)	.001
Participants performing $\geq 2$ measurements per day, n (%)	21 (91)	23 (100)	11 (65)	.008

<sup>a</sup>Obtained using a *t* test or 2-proportion *z* test.

**Figure 2.** Overview of FibriCheck results comparing the proportion of each rhythm diagnosis between the first and second month of follow-up, with significance determined using 2-proportion z tests. AF: atrial fibrillation; PVC: premature ventricular contraction, SR: sinus rhythm.



On average, 2.7 (SD 0.9) measurements were performed per day. Compliance was significantly better in the 1st vs 2nd month (Table 2), and 21 participants (21/23, 91%) indicated that performing 3 measurements per day was manageable. Patients attributed the decrease in compliance mostly to feeling better and returning to work.

### Participants and Primary Care Physician's Experience

Most participants found that using FibriCheck after discharge was reassuring and made them feel safer (17/23, 74%). Two of the participants experienced mild stress, and 1 participant experienced severe distress. Even though some participants experienced stress while performing the measurements, 20 out of 23 participants (87%) recommended it to be used in common practice. The surveyed PCPs experienced positive value by the added use of the app as a follow-up method as they felt more involved in the postoperative rehabilitation.

## Discussion

PPG technology-based smartphone apps show great promise to improve follow-up during rehabilitation after cardiac surgery. To our knowledge, this study is the first to evaluate the

systematic use of a PPG-based app for remote heart rhythm monitoring in the first 1-2 months after cardiac surgery. In 23 patients with 30 days of follow-up, AF was detected in 4 (17%) patients. In 2 of these patients, the detection of AF by FibriCheck led to the diagnosis of an underlying complication. This finding suggests that late-onset or recurrent POAF after cardiac surgery may be an alarm symptom. In 1 patient, a late-onset AV block Mobitz II was detected as the irregular rhythm triggered a false positive registration of atrial flutter.

### Study Population

The population of our study consisted of a mixed cohort of surgical patients, representative for a tertiary care center. The mean age and CHA<sub>2</sub>DS<sub>2</sub>-VASc score in our population were lower compared to other short-term, invasive and noninvasive follow-up studies [14,20]. While this may indicate that our population was at a relatively low risk for AF, POAF occurred in 9 out of 23 patients (39%) during hospitalization [7]. It would be incorrect to draw conclusions from the demographic variables and compare these to the risk factors reported in the literature because of the limited number of included patients.

## Detection of Events by FibriCheck

Even with a limited sample size, our study showed an incidence of POAF after discharge of 17% (4 out of 23 participants). The value of FibriCheck is not limited to the detection of AF as it may enable early diagnosis of postoperative complications. While FibriCheck is not (yet) able to specifically detect atrioventricular conduction abnormalities, the app enabled the detection of a second-degree AV block by giving a false positive registration for atrial flutter [21]. Furthermore, this case indicates the importance of confirming the diagnosis of arrhythmias detected by PPG with a 12-lead ECG. In 1 patient in our study, POAF recurrence was most likely caused by a pericardial tamponade [22]. Likewise, we detected an episode of late-onset AF related to pneumonia. These associations suggest that AF may serve as a marker for underlying complications, supported by the established relation between AF and inflammation. In the case of pericardial tamponade or pneumonia, it is likely that a combination of hemodynamic effects and inflammatory status led to AF [23].

When comparing the first month of follow-up to the second month, several significant differences were noted. The proportion of normal or sinus rhythm measurements was significantly higher in the second month, and consequently, the numbers of warning and urgent measurements were significantly lower. This is in keeping with evidence that most recurrence arises within the first month after surgery [9,14,24].

## Strengths and Weaknesses of the App

The ability to perform measurements at any moment is one of the most attractive features of PPG-based smartphone technology, especially in a clinical culture in which remote follow-up is becoming increasingly important. It also offers a unique opportunity to study and broaden knowledge of the occurrence of postcardiac surgery complications as we observed with our patients with AV block, pericardial tamponade, and pneumonia. In our experience, FibriCheck accurately categorizes the urgency of a measurement, making it easier and less time-consuming for researchers or physicians to use the app.

Nonetheless, the implementation of PPG-based apps still faces several obstacles. First, certain arrhythmias, such as an AV block, cannot be diagnosed as such. Second, review time can cause some inconvenience, so active follow-up is warranted. When considering routine implementation of a PPG-based app into a postoperative setting, dedicated personnel is likely needed to ensure a streamlined follow-up and lower the additional workload. Basic understanding of PPG has proven useful to interpret the measurements ourselves. Finally, patient compliance, with sufficient measurements performed per day, seems essential to the success of smartphone-based PPG technology in the diagnosis and follow-up of AF. In this study,

only 7 out of 23 participants (30%) performed the required 3 measurements per day, yet 21 out of 23 participants (91%) performed 2 measurements per day. As compliance decreased over time in this study, built-in reminders may ensure sufficient measurements. On the other hand, all of the arrhythmias with clinical implications diagnosed in our study were detected because patients performed a PPG registration because they were symptomatic. This highlights the importance of having a diagnostic tool readily available for patients at risk for arrhythmias, in the postoperative setting or otherwise.

## Study Limitations and Implications for Future Studies

When designing this study, it was anticipated that a large proportion of the elderly population, with a higher risk of arrhythmias, would not possess a smartphone [25] and thus be excluded from participation, thereby creating a selection bias. Indeed, the average age of included patients was rather young (mean 60.2, SD 12 years) and significantly younger than the patients excluded due to the lack of smartphone possession (mean 77, SD 8.5 years,  $P < .001$ ). As the fraction of adults aged 65 years or older using smartphones rose from 18% to 42% between 2013 and 2016, the implications for future studies may be diminished in the future [25]. Alternatively, future projects can actively reduce this bias by providing devices or involving caregivers.

While the small sample size of this study prevents us from offering new insights on arrhythmia occurrence after cardiac surgery, our study did find an interesting association between POAF and underlying complications. The COVID-19 outbreak (in March 2020) hindered patient inclusion and follow-up consultations. Future trials will have a larger study population and include a scaled questionnaire of patient and PCP experience to enable continued improvement of the application.

## Conclusions

This study was the first to evaluate FibriCheck in a setting after cardiac surgery. Implementation of smartphone-based PPG technology enabled the detection of AF and other (rhythm-related) complications. After discharge, even with only 23 patients included, FibriCheck detected POAF in 4 patients (17%) and a second-degree AV block in 1 patient. An association between AF detection and an underlying complication was found in 2 patients. Early detection of these complications by FibriCheck likely improved the patients' clinical outcomes. Therefore, smartphone-based PPG technology may supplement rehabilitation after cardiac surgery by acting as a sentinel for underlying complications, rhythm-related or otherwise. With dedicated personnel and a streamlined workflow for the management of detected arrhythmias, the systemic implementation of FibriCheck into follow-up after cardiac surgery is very promising.

## Acknowledgments

For the purpose of this study, Qompium NV provided support by allowing use of the photoplethysmography-based smartphone app "FibriCheck" without cost to the research team or the patients in this study.

## Authors' Contributions

ML, LH, FR, PH, TP, and LG contributed to the study design. ML included patients. ML and LH ensured follow-up of measurement results. ML performed the data collection and analysis. All authors contributed to the manuscript and approved the final version to be published.

## Conflicts of Interest

LG and TP are employees of Qompium NV, the developer of the specific photoplethysmography app "FibriCheck" used in this study. LG is CEO and shareholder of Qompium NV.

## Multimedia Appendix 1

An example of a normal PPG measurement in the dashboard of the application.

[[PDF File \(Adobe PDF File\), 337 KB - mhealth\\_v9i4e26519\\_app1.pdf](#)]

## References

1. Clay-Weinfeld K, Callans M. Common Postcardiothoracic Surgery Arrhythmias. *Crit Care Nurs Clin North Am* 2019 Oct;31(3):367-388. [doi: [10.1016/j.cnc.2019.05.006](https://doi.org/10.1016/j.cnc.2019.05.006)] [Medline: [31351556](https://pubmed.ncbi.nlm.nih.gov/31351556/)]
2. Maesen B, Nijs J, Maessen J, Allessie M, Schotten U. Post-operative atrial fibrillation: a maze of mechanisms. *Europace* 2012 Mar 06;14(2):159-174 [FREE Full text] [doi: [10.1093/europace/eur208](https://doi.org/10.1093/europace/eur208)] [Medline: [21821851](https://pubmed.ncbi.nlm.nih.gov/21821851/)]
3. Hernández-Leiva E, Alvarado P, Dennis RJ. Postoperative Atrial Fibrillation: Evaluation of its Economic Impact on the Costs of Cardiac Surgery. *Braz J Cardiovasc Surg* 2019;34(2):179-186 [FREE Full text] [doi: [10.21470/1678-9741-2018-0218](https://doi.org/10.21470/1678-9741-2018-0218)] [Medline: [30916128](https://pubmed.ncbi.nlm.nih.gov/30916128/)]
4. Gudbjartsson T, Helgadóttir S, Sigurdsson MI, Taha A, Jeppsson A, Christensen TD, et al. New-onset postoperative atrial fibrillation after heart surgery. *Acta Anaesthesiol Scand* 2020 Feb 02;64(2):145-155. [doi: [10.1111/aas.13507](https://doi.org/10.1111/aas.13507)] [Medline: [31724159](https://pubmed.ncbi.nlm.nih.gov/31724159/)]
5. Lomivorotov VV, Efremov SM, Pokushalov EA, Karaskov AM. New-Onset Atrial Fibrillation After Cardiac Surgery: Pathophysiology, Prophylaxis, and Treatment. *J Cardiothorac Vasc Anesth* 2016 Jan;30(1):200-216. [doi: [10.1053/j.jvca.2015.08.003](https://doi.org/10.1053/j.jvca.2015.08.003)] [Medline: [26597469](https://pubmed.ncbi.nlm.nih.gov/26597469/)]
6. Peretto G, Durante A, Limite L, Cianflone D. Postoperative arrhythmias after cardiac surgery: incidence, risk factors, and therapeutic management. *Cardiol Res Pract* 2014;2014:615987 [FREE Full text] [doi: [10.1155/2014/615987](https://doi.org/10.1155/2014/615987)] [Medline: [24511410](https://pubmed.ncbi.nlm.nih.gov/24511410/)]
7. Greenberg J, Lancaster T, Schuessler R, Melby S. Postoperative atrial fibrillation following cardiac surgery: a persistent complication. *Eur J Cardiothorac Surg* 2017 Oct 01;52(4):665-672. [doi: [10.1093/ejcts/ezx039](https://doi.org/10.1093/ejcts/ezx039)] [Medline: [28369234](https://pubmed.ncbi.nlm.nih.gov/28369234/)]
8. Kawamoto N, Fujita T, Fukushima S, Shimahara Y, Kume Y, Matsumoto Y, et al. Late onset of atrial fibrillation in patients undergoing mitral valve repair for type II dysfunction. *J Cardiol* 2018 Apr;71(4):346-351 [FREE Full text] [doi: [10.1016/j.jicc.2017.11.010](https://doi.org/10.1016/j.jicc.2017.11.010)] [Medline: [29275952](https://pubmed.ncbi.nlm.nih.gov/29275952/)]
9. Magruder J, Collica S, Belmustakov S, Crawford TC, Grimm JC, Cameron DE, et al. Predictors of Late-Onset Atrial Fibrillation Following Isolated Mitral Valve Repairs in Patients With Preserved Ejection Fraction. *J Card Surg* 2016 Aug 15;31(8):486-492. [doi: [10.1111/jocs.12774](https://doi.org/10.1111/jocs.12774)] [Medline: [27302368](https://pubmed.ncbi.nlm.nih.gov/27302368/)]
10. Melduni RM, Schaff HV, Bailey KR, Cha SS, Ammash NM, Seward JB, et al. Implications of new-onset atrial fibrillation after cardiac surgery on long-term prognosis: a community-based study. *Am Heart J* 2015 Oct;170(4):659-668 [FREE Full text] [doi: [10.1016/j.ahj.2015.06.015](https://doi.org/10.1016/j.ahj.2015.06.015)] [Medline: [26386789](https://pubmed.ncbi.nlm.nih.gov/26386789/)]
11. Konstantino Y, Zelnik Yovel D, Friger MD, Sahar G, Knyazer B, Amit G. Postoperative Atrial Fibrillation Following Coronary Artery Bypass Graft Surgery Predicts Long-Term Atrial Fibrillation and Stroke. *Isr Med Assoc J* 2016 Dec;18(12):744-748 [FREE Full text] [Medline: [28457078](https://pubmed.ncbi.nlm.nih.gov/28457078/)]
12. Ambrosetti M, Tramarin R, Griffo R, De Feo S, Fattiroli F, Vestri A, ISYDEICAROS Investigators of the Italian Society for Cardiovascular Prevention, Rehabilitation and Epidemiology (IACPR-GICR). Late postoperative atrial fibrillation after cardiac surgery: a national survey within the cardiac rehabilitation setting. *J Cardiovasc Med (Hagerstown)* 2011 Jul;12(6):390-395. [doi: [10.2459/JCM.0b013e328346a6d3](https://doi.org/10.2459/JCM.0b013e328346a6d3)] [Medline: [21508846](https://pubmed.ncbi.nlm.nih.gov/21508846/)]
13. Cioffi G, Mureddu G, Cemin C, Russo TE, Pellegrini A, Terrasi F, et al. Characterization of post-discharge atrial fibrillation following open-heart surgery in uncomplicated patients referred to an early rehabilitation program. *Ital Heart J* 2001 Jul;2(7):519-528. [Medline: [11501961](https://pubmed.ncbi.nlm.nih.gov/11501961/)]
14. Lowres N, Mulcahy G, Jin K, Gallagher R, Neubeck L, Freedman B. Incidence of postoperative atrial fibrillation recurrence in patients discharged in sinus rhythm after cardiac surgery: a systematic review and meta-analysis. *Interact Cardiovasc Thorac Surg* 2018 Mar 01;26(3):504-511. [doi: [10.1093/icvts/ivx348](https://doi.org/10.1093/icvts/ivx348)] [Medline: [29161419](https://pubmed.ncbi.nlm.nih.gov/29161419/)]
15. El-Chami MF, Merchant FM, Smith P, Levy M, Nelms AG, Brand J, et al. Management of New-Onset Postoperative Atrial Fibrillation Utilizing Insertable Cardiac Monitor Technology to Observe Recurrence of AF (MONITOR-AF). *Pacing Clin Electrophysiol* 2016 Oct 22;39(10):1083-1089. [doi: [10.1111/pace.12949](https://doi.org/10.1111/pace.12949)] [Medline: [27570120](https://pubmed.ncbi.nlm.nih.gov/27570120/)]

16. Li KHC, White FA, Tipoe T, Liu T, Wong MC, Jesuthasan A, et al. The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review. *JMIR Mhealth Uhealth* 2019 Feb 15;7(2):e11606 [FREE Full text] [doi: [10.2196/11606](https://doi.org/10.2196/11606)] [Medline: [30767904](https://pubmed.ncbi.nlm.nih.gov/30767904/)]
17. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile Phone-Based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: Diagnostic Accuracy Study of the FibriCheck App. *JMIR Mhealth Uhealth* 2019 Mar 27;7(3):e12284 [FREE Full text] [doi: [10.2196/12284](https://doi.org/10.2196/12284)] [Medline: [30916656](https://pubmed.ncbi.nlm.nih.gov/30916656/)]
18. Verbrugge F, Proesmans T, Vijgen J, Mullens W, Rivero-Ayerza M, Van Herendael H, et al. Atrial fibrillation screening with photo-plethysmography through a smartphone camera. *Europace* 2019 Aug 01;21(8):1167-1175. [doi: [10.1093/europace/euz119](https://doi.org/10.1093/europace/euz119)] [Medline: [31056678](https://pubmed.ncbi.nlm.nih.gov/31056678/)]
19. Hermans A, van der Velden RMJ, Gawalko M, Verhaert D, Desteghe L, Duncker D, TeleCheck-AF investigators. On-demand mobile health infrastructures to allow comprehensive remote atrial fibrillation and risk factor management through teleconsultation. *Clin Cardiol* 2020 Dec;43(11):1232-1239 [FREE Full text] [doi: [10.1002/clc.23469](https://doi.org/10.1002/clc.23469)] [Medline: [33030259](https://pubmed.ncbi.nlm.nih.gov/33030259/)]
20. Lowres N, Mulcahy G, Gallagher R, Ben Freedman S, Marshman D, Kirkness A, et al. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. *Eur J Cardiothorac Surg* 2016 Jul;50(1):44-51. [doi: [10.1093/ejcts/ezv486](https://doi.org/10.1093/ejcts/ezv486)] [Medline: [26850266](https://pubmed.ncbi.nlm.nih.gov/26850266/)]
21. Lamberigts M, Van Hoof L, Haemers P, Herijgers P, Proesmans T, Grieten L. Detection of a 2nd degree AV block by a photoplethysmography based smartphone application after recent mitral valve surgery. 2020 Presented at: ESC Congress 2020; August 30, 2020; Amsterdam, Netherlands. [doi: [10.2196/26519](https://doi.org/10.2196/26519)]
22. Haskell RJ, French WJ. Cardiac tamponade as the initial presentation of malignancy. *Chest* 1985 Jul;88(1):70-73. [doi: [10.1378/chest.88.1.70](https://doi.org/10.1378/chest.88.1.70)] [Medline: [4006558](https://pubmed.ncbi.nlm.nih.gov/4006558/)]
23. Harada M, Van Wagoner DR, Nattel S. Role of inflammation in atrial fibrillation pathophysiology and management. *Circ J* 2015;79(3):495-502 [FREE Full text] [doi: [10.1253/circj.CJ-15-0138](https://doi.org/10.1253/circj.CJ-15-0138)] [Medline: [25746525](https://pubmed.ncbi.nlm.nih.gov/25746525/)]
24. Cioffi G, Cemin C, Russo TE, Pellegrini A, Terrasi F, Ferrario G. Post-discharge recurrences of new-onset atrial fibrillation following cardiac surgery: impact of low-dose amiodarone and beta-blocker prophylaxis. *Ital Heart J* 2000 Oct;1(10):691-697. [Medline: [11061366](https://pubmed.ncbi.nlm.nih.gov/11061366/)]
25. Anderson MA, Perrin A. Tech Adoption Climbs Among Older Adults. Pew Research Center. 2017 May 17. URL: [https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2017/05/PI\\_2017.05.17\\_Older-Americans-Tech\\_FINAL.pdf](https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2017/05/PI_2017.05.17_Older-Americans-Tech_FINAL.pdf) [accessed 2020-10-01]

## Abbreviations

**AF:** atrial fibrillation

**AV:** atrioventricular

**CHA<sub>2</sub>DS<sub>2</sub>-VASc:** Congestive heart failure, Hypertension, Age (>75y), Diabetes, Stroke/transient ischemic attack, Vascular disease, previous myocardial infarction, Age (65-75y) and Sex category

**ECG:** electrocardiogram

**PCP:** primary care physician

**POAF:** postoperative atrial fibrillation

**PPG:** photoplethysmography

**UZ:** University Hospitals

*Edited by L Buis; submitted 16.12.20; peer-reviewed by K Deldar, S Smith; comments to author 16.01.21; revised version received 28.01.21; accepted 24.03.21; published 15.04.21.*

*Please cite as:*

*Lamberigts M, Van Hoof L, Proesmans T, Vandervoort P, Grieten L, Haemers P, Rega F*

*Remote Heart Rhythm Monitoring by Photoplethysmography-Based Smartphone Technology After Cardiac Surgery: Prospective Observational Study*

*JMIR Mhealth Uhealth* 2021;9(4):e26519

URL: <https://mhealth.jmir.org/2021/4/e26519>

doi: [10.2196/26519](https://doi.org/10.2196/26519)

PMID: [33856357](https://pubmed.ncbi.nlm.nih.gov/33856357/)

©Marie Lamberigts, Lucas Van Hoof, Tine Proesmans, Pieter Vandervoort, Lars Grieten, Peter Haemers, Filip Rega. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 15.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Perception of Purposeful and Recreational Smartphone Use in Physiotherapy: Randomized Controlled Trial

Martina Bientzle<sup>1</sup>, Dr rer nat; Anne Restle<sup>2</sup>, BSc; Joachim Kimmerle<sup>1</sup>, Dr rer nat

<sup>1</sup>Leibniz-Institut für Wissensmedien, Tuebingen, Germany

<sup>2</sup>PT Akademie Tuebingen, Tuebingen, Germany

**Corresponding Author:**

Martina Bientzle, Dr rer nat

Leibniz-Institut für Wissensmedien

Schleichstr. 6

Tuebingen, 72076

Germany

Phone: 49 7071 979 120

Fax: 49 7071 979 105

Email: [m.bientzle@iwm-tuebingen.de](mailto:m.bientzle@iwm-tuebingen.de)

## Abstract

**Background:** Many people constantly use their smartphones in all kinds of situations. Often smartphones are used in a meaningful and targeted way, but frequently they are used as a pastime without any purpose. This also applies to patients and therapists in treatment situations.

**Objective:** The aim of this study was to investigate how purposeful smartphone use compared with recreational smartphone use (by a physiotherapist or by a patient) influenced the perception of a physiotherapeutic treatment situation. We examined the impact of smartphone use during a physiotherapy session on the perception of the physiotherapist, evaluation of attentiveness, and evaluation of smartphone use in physiotherapy in general.

**Methods:** Members of various music and sports clubs were invited to participate in an online randomized controlled trial. Participants were randomly assigned to one of four conditions. They watched a video in which a physiotherapeutic treatment was shown and in which a smartphone was used or not used in the following four different ways: (1) therapeutically purposeful use, (2) recreational use by the physiotherapist (looking at the phone from time to time with no therapeutic purpose), (3) recreational use by the patient, and (4) no smartphone use (control condition). After watching the video, the participants indicated their perception of the physiotherapist's professional competence, social competence, and empathetic behavior. They also rated the physiotherapist's and patient's attentiveness and evaluated the usage of smartphones generally in physiotherapy.

**Results:** The analysis included 118 participants (63 women and 55 men). When the physiotherapist used the smartphone in a purposeful way, the physiotherapist was perceived as more professionally competent ( $P=.007$ ), socially competent ( $P=.03$ ), and empathetic ( $P=.04$ ) than if the physiotherapist used it with no therapeutic purpose. These effects occurred because recreational smartphone use by the physiotherapist was evaluated more negatively than the behavior in the control condition (professional competence:  $P=.001$ ; social competence:  $P=.03$ ; empathy:  $P=.04$ ). Moreover, when the physiotherapist used the smartphone in a recreational way, the physiotherapist was perceived as being less attentive ( $P<.001$ ). Likewise, when the patient used the smartphone in a recreational way, the patient was perceived as being less attentive ( $P<.001$ ). Finally, smartphone use in physiotherapy was rated as more positive in general when the smartphone was used in a purposeful way compared with the conditions in which the physiotherapist or patient looked at the smartphone with no therapeutic purpose ( $P<.001$ ). This positive evaluation occurred because purposeful use led to a more positive rating than no smartphone use ( $P<.001$ ,  $R=0.42$ ).

**Conclusions:** Smartphones are only appropriate for therapists and patients if they are used directly for a therapeutic purpose. Otherwise, it is better not to use smartphones during treatment.

**Trial Registration:** AsPredicted (aspredicted.org) #24740; <https://aspredicted.org/blind.php?x=vv532i>

(*JMIR Mhealth Uhealth* 2021;9(4):e25717) doi:[10.2196/25717](https://doi.org/10.2196/25717)

**KEYWORDS**

smartphone use; phubbing; physiotherapy; smartphone; therapy; patients; therapists; therapeutic; treatment

## Introduction

### Background

Smartphones and other mobile devices are ubiquitous. They enable their users to complete a large number of tasks with little effort. They can be used not only for all kinds of everyday tasks, but also expressly for medical diagnostics [1-4] and therapeutic purposes [5-7]. Whether the use of smartphones is perceived as meaningful and helpful, or rather as unnecessary and even disruptive, certainly depends on the concrete way in which they are used. On the one hand, smartphones are versatile tools that can be used for a variety of tasks and activities like social interactions [8,9], education [10-12], and work-related activities [13,14]. On the other hand, many people are often distracted by their smartphones as they constantly look at them and check whether there is something new, even if this has nothing to do with the actual situation they are in at that moment [15,16]. Despite the positive possibilities offered by smartphones, the devices have dangers. People tend to spend several hours a day with their devices and social interactions can be disturbed [15]. Communication via the smartphone has become faster, easier, and a matter of course. The expectation of being “permanently online and permanently connected” seems to be established in society, in private life, and at work [17-19]. This is particularly noticeable in certain behaviors, such as the frequent habit of briefly “checking” the phone [20-22]. People focusing on their smartphones instead of their physically present communication partners is such a common situation that the term “phubbing” has been created to describe this phenomenon [23-25].

Checking behavior refers to people’s habit of constantly inspecting their smartphones. This includes receiving messages from other people and reports of news through websites or social media. Habits are repetitive procedures and activities in certain situations. This checking habit is a kind of ritual for many people, which mostly happens unconsciously [25,26]. The short repetitive checks are thus automatic behaviors and can be increased by external stimuli, such as visual and auditory stimuli [26]. Constant distraction and interruption can lead to errors and a reduction in efficiency [27]. Negative effects may occur owing to minor interruptions in everyday activities, such as working, learning, and driving. In some cases, smartphone use can even be considered an addiction [28-30]. In everyday life, smartphones accompany many patients and therapists in physiotherapy sessions. As in other situations, smartphones can offer opportunities for therapy [31-33], but at the same time, they can present risks.

### Smartphone Use in Physiotherapy

In physiotherapy, new exercises are often shown to build up muscles or train balance and body awareness. One technique that is often used to support such learning processes is learning through observing one’s own behavior or the behaviors of human role models [34-37]. This includes learning by observing a demonstration of a target movement. Video exercise instruction [38-40], video feedback [41], and video self-modeling [42,43] are possible implementations of this principle.

Through video self-modeling (ie, a video recording of the patient herself/himself during the implementation of a motor movement sequence), visual evidence is produced. This differs from the video exercise instruction, which shows a film excerpt with the motion sequence being performed by an expert or professional. Many studies have confirmed the usefulness of video feedback to improve athletic performance [44,45]. In physiotherapy, a video recording can be used to alert the patient to errors and to request correction of these errors. With the help of smartphones, video recording is made easier and is possible for any patient. Therefore, the video feedback approach with modern smartphones in physiotherapy should be considered as a possible support method.

The interaction between therapists and patients is a complex construct. The therapeutic relationship between therapists and patients is an important factor for the success of physiotherapy. In this type of relationship, aspects, such as social and professional competence and empathy, are particularly important [46-48]. To support successful treatment, attention and motivation are relevant. Distractions of any kind, for example, by a smartphone, can have a negative effect on the success of therapy. In addition, the perceived use of smartphones can influence the therapy. Professional and social competence as well as empathy and attentiveness are factors that contribute to the success of physiotherapeutic treatments. At the same time, these are variables that are potentially affected by the use of smartphones, and accordingly, represent relevant outcome variables for this study [49,50]. A recent study showed that observing the use of smartphones in social interactions influences the perception of smartphone users’ warmth and competence [51]. So far, it is unclear whether this also applies to physiotherapeutic treatment situations and whether the way smartphones are used (purposeful smartphone use vs smartphone use in the sense of a checking habit) influences the perception of the treatment situation and the therapist.

### Outcome Variables and Hypotheses

In order to ensure a positive effect of the therapeutic relationship, *perceived professional competence* of the physiotherapist is central [46,52-54]. The use of a smartphone may influence the perception of the physiotherapist’s professional competence. Targeted and purposeful smartphone use can have a positive effect on the perception of professional competence in cases where use of a smartphone obviously supports therapeutic treatment. On the other hand, smartphone use in the sense of a checking habit can have a negative effect on perceived professional competence, as the device distracts from the treatment and disturbs its progression. On the basis of these considerations, we state the following hypothesis (hypothesis 1): The way the smartphone is used has an impact on participants’ perceived professional competence of physiotherapists. Physiotherapists’ professional competence will be perceived as more pronounced if they use the smartphone in a purposeful way than if they look at it from time to time with no therapeutic purpose.

In addition to professional competence, *perceived social competence* of the physiotherapist is another relevant factor influencing the success of treatment [46,54]. Social competence

refers to physiotherapists' interpersonal and communicative skills. This aspect is important for interpersonal communication and relationship building. A change in the perception of physiotherapists' social competence with regard to the way they use smartphones is to be expected. Purposeful use of smartphones can have a positive effect on the perception of social competence, when the smartphone serves as an aid during treatment. Conversely, smartphone use in the sense of a checking habit can have a negative effect on perceived social competence. Based on these considerations, we state the following hypothesis (hypothesis 2): The way the smartphone is used has an impact on participants' perceived social competence of physiotherapists. Physiotherapists' social competence will be perceived as more pronounced if they use the smartphone in a purposeful way than if they look at it from time to time with no therapeutic purpose.

Adequate empathetic behavior by therapists is also important to promote a positive therapeutic relationship and treatment success. Through *perceived empathetic behavior*, patients feel understood and build a positive relationship and trust with the treating therapist [54-57]. With respect to perceived empathetic behavior, we state the following hypothesis (hypothesis 3): The way the smartphone is used has an impact on participants' perceived empathetic behavior of physiotherapists. Physiotherapists' empathetic behavior will be perceived as more pronounced if they use the smartphone in a purposeful way than if they look at it from time to time with no therapeutic purpose.

As an open research question, we examined whether the assumed differences in perceived professional competence, perceived social competence, and perceived empathetic behavior occur owing to the fact that (1) purposeful use of the smartphone leads to a *higher* rating of perceived professional competence, perceived social competence, and perceived empathetic behavior in comparison with no smartphone use or (2) the checking behavior of the physiotherapist leads to a *lower* rating in comparison with no smartphone use.

In order to be able to ensure successful therapy, a good working atmosphere must be created. One factor that can positively influence the working atmosphere is attention. It is important not to be distracted in order to be able to focus one's full attention on something [58]. Using a smartphone to answer messages, etc, at the same time that an exercise in physiotherapy is being explained can lead to attention problems and consequently to performance degradation [59]. The reason for this is that the attention capacity of a person is not sufficient to perform both tasks at the same time [60,61]. Smartphone use immediately directs attention to the stimuli of the smartphone, disrupting the attention that is needed for the physiotherapy exercise. Thus, we expected the way the smartphone is used to have an impact on how attentively physiotherapists and patients are perceived by the participants. Thus, we state the following hypotheses: hypothesis 4a, physiotherapists will be perceived as being less attentive when they look at the smartphone from time to time with no therapeutic purpose; and hypothesis 4b,

patients will be perceived as being less attentive when they look at the smartphone from time to time with no therapeutic purpose.

In this context, it is also relevant to investigate how the concrete use of smartphones in the physiotherapeutic treatment situation affects how the participants evaluate the use of smartphones in general. We state the following hypothesis (hypothesis 5): The way the smartphone is used has an impact on participants' evaluation of smartphone use in physiotherapy in general. Smartphone use in physiotherapy will be rated as more positive if the smartphone is used in a purposeful way than if physiotherapists or patients look at it from time to time with no therapeutic purpose.

As an open research question, we examined whether the assumed differences are owing to the fact that (1) purposeful use of the smartphone leads to a *more* positive rating in comparison with no smartphone use or (2) use of the smartphone with no therapeutic purpose leads to a *less* positive rating in comparison with no smartphone use.

## Methods


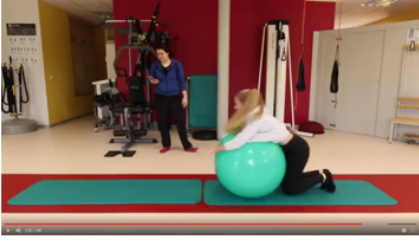

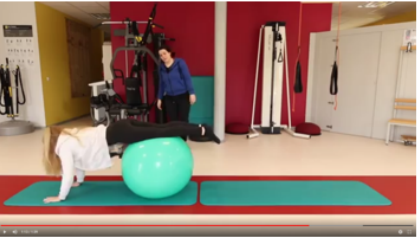
### Ethical Approval

This study was approved by the Ethics Committee of the Leibniz-Institut fuer Wissensmedien (Tuebingen, Germany; approval number: LEK 2019/025).

### Study Design

The data for this study were collected in an online survey. The survey contained several questionnaires and a video that differed depending on the experimental condition. We used a video presentation, because this allowed for standardized manipulation of the conditions. The video clips lasted about 2 minutes. They showed a physiotherapeutic treatment with different types of smartphone use, involving a female physiotherapist and a female patient, who were both portrayed by actresses. The first and the last scenes of the video were identical in all of the conditions. First, the patient was greeted and asked about the current back pain. In the following part, the physiotherapist provided instructions for an exercise with a Pezzi ball. Here, the physiotherapist demonstrated the exercise and the patient imitated it. In this part, the conditions differed only with regard to smartphone use (Figure 1). The conditions were as follows: (1) *Purposeful use*, the smartphone was used to support the physiotherapy, in the sense that the patient was filmed with a smartphone doing an exercise to provide qualified feedback; (2) *Check therapist*, the physiotherapist looked at the smartphone from time to time with no therapeutic purpose; (3) *Check patient*, the patient looked at the smartphone from time to time with no therapeutic purpose; (4) *Control*, no smartphone was used during the exercise. At the end of the video, the patient was asked to do the exercise again. This scene was identical in all conditions. The participants were randomly assigned to watch one of the four videos.

Figure 1. Experimental conditions.

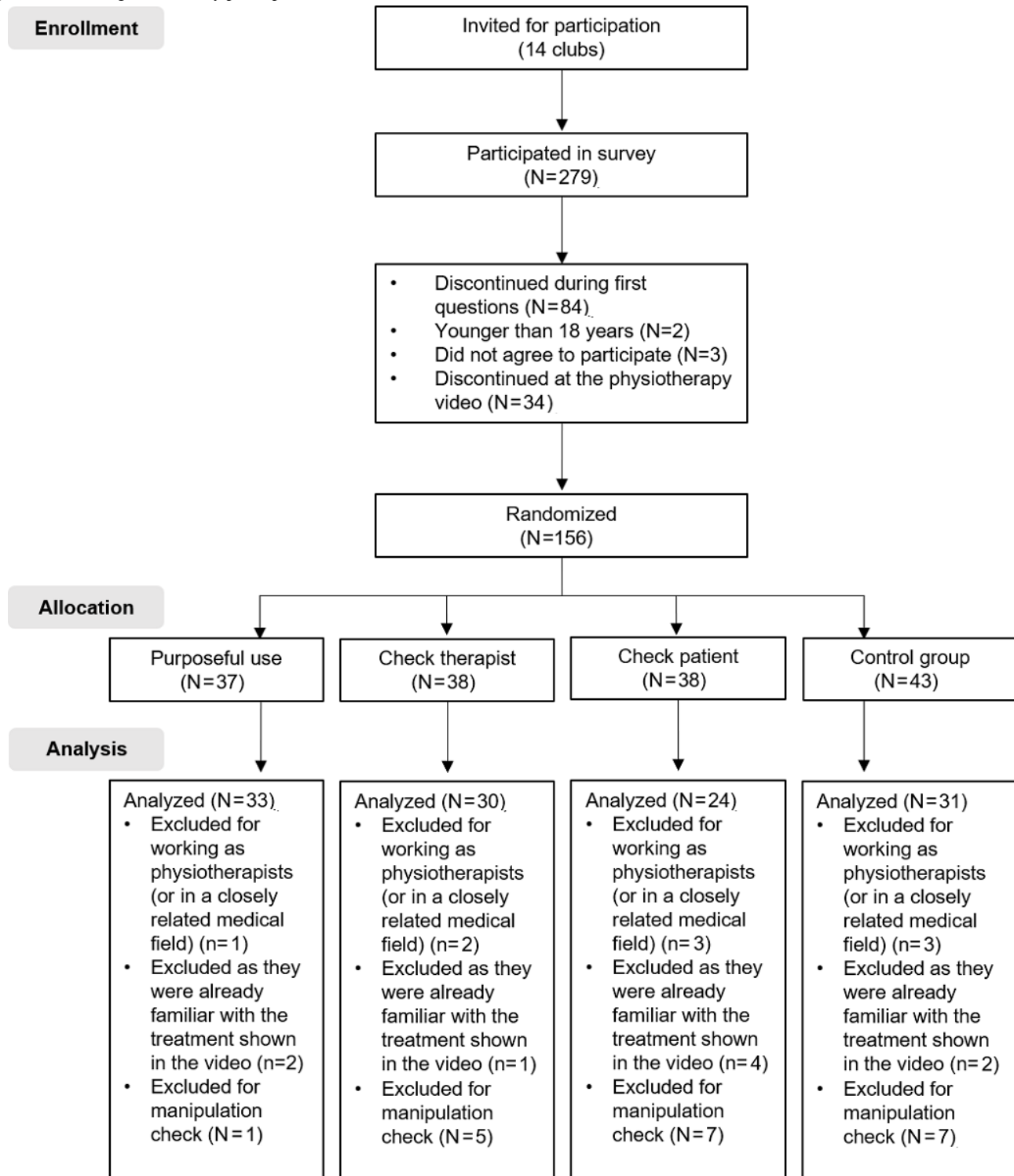
Condition	Duration	Smartphone use	Screenshot of the video
Purposeful use	2:09	A smartphone could be seen in 4 scenes of the video: The physiotherapist filmed the patient with the smartphone; physiotherapist and patient watched the video together on the smartphone; in two other scenes, the smartphone was present, but nobody looked at it.	
Check therapist	1:48	A smartphone could be seen in 4 scenes of the video: The physiotherapist checked the smartphone 4 times during the video.	
Check patient	1:47	A smartphone could be seen in 4 scenes of the video: The patient checked the smartphone 4 times during the video.	
Control group	1:38	No smartphone was present in the video.	

## Participants

A power analysis for analysis of variance (ANOVA) with  $\alpha=.05$ , an intended power of 95%, and a large effect size of 0.40 revealed a required sample size of 112. Members of various music and sports clubs in Germany (Federal State of Baden-Wuerttemberg) were invited via the clubs' internal email distribution lists to participate from June to July 2019 in an online study. This was a relevant sample as both athletes and musicians often need physiotherapy [62]. The invitation email provided basic information about the study. This comprised the inclusion criteria (good German language skills and a minimum age of 18 years) and the formal requirements to participate in the study (the required technical device, ie, a computer, laptop, or smartphone with speakers or headphones, and the expected

time required). A total of 279 people from 14 clubs took part in the survey. After applying the predefined exclusion criteria (Figure 2), we included 118 participants in the analysis. The age of the participants was between 18 and 74 years (mean 39.83, SD 15.15 years). Of the 118 participants, 63 stated they were female and 55 stated they were male. Additionally, 75 participants were employed or self-employed, 19 were students, three were in vocational training, eight were retirees, and 14 indicated other occupations, such as public officials. Seven participants did not answer the occupation question (multiple answers were possible). Most participants ( $n=107$ , 86.4%) indicated that they use their smartphones within the first hour after getting up, and more than half of the participants ( $n=61$ , 51.7%) indicated that they use their smartphones between 1 and 4 hours a day.

Figure 2. Flow diagram for study participants.



## Procedure

Initially, the participants provided information about their age, gender, current occupation, and daily smartphone usage. Thereafter, they were asked to imagine that they were undergoing physiotherapy for back pain. After that, the participants were randomly assigned to one of the four experimental conditions equally and they watched the appropriate video. Randomization was carried out by using Qualtrics software (Qualtrics). All participants were blinded to group allocation.

After watching the video, we conducted a manipulation check to ensure that the participants recognized the experimental treatment. They then indicated their perception of the professional competence, social competence, and empathetic behavior of the physiotherapist. They also rated the physiotherapist's and patient's attentiveness and evaluated the usage of smartphones in physiotherapy in general. Finally, the participants were debriefed. They were given the investigators' contact information, and a link was provided to enable them to take part in a raffle for Amazon vouchers (two vouchers of €50

(USD 60) each and five vouchers of €20 (USD 23) each. Participation in the study took about 20 to 30 minutes.

## Measures

The online study was created using the Qualtrics software. The questionnaire of Willson and McNamara was used to measure *professional competence* and *social competence* [63]. This

questionnaire compares eight pairs of adjectives to measure professional competence and nine pairs of adjectives to measure social competence on 9-point scales. These items are shown in [Textbox 1](#). The reliability of the scales was determined by calculating the Cronbach alpha coefficient. There were good internal consistencies for the professional competence ( $\alpha=.894$ ) and social competence scales ( $\alpha=.897$ ).

**Textbox 1.** Measurement of perceived professional and social competence. \*Reverse-coded items

<p><b>Professional competence</b></p> <p>Unprofessional (1) to professional (9)</p> <p>Experienced (1) to inexperienced (9) *</p> <p>Not thorough (1) to thorough (9)</p> <p>Careful (1) to careless (9) *</p> <p>Incompetent (1) to competent (9)</p> <p>Trained (1) to untrained (9) *</p> <p>Not appealing (1) to appealing (9)</p> <p>Confident (1) to unconfident (9) *</p> <p><b>Social competence</b></p> <p>Friendly (1) to unfriendly (9) *</p> <p>Impolite (1) to polite (9)</p> <p>Attentive (1) to not attentive (9) *</p> <p>Unkind (1) to kind (9)</p> <p>Pleasant (1) to unpleasant (9) *</p> <p>Not nice (1) to nice (9)</p> <p>Caring (1) to not caring (9) *</p> <p>Insensitive (1) to sensitive (9)</p> <p>Sympathetic (1) to unsympathetic (9) *</p>
--

*Empathetic behavior* of the physiotherapist was captured with a modified and shortened version of the CARE scale [64]. We modified the scale to fit the physiotherapeutic situation (rating the physiotherapist, not a physician). We used five items that

participants rated on a 5-point Likert scale. Internal consistency for the scale was good ( $\alpha=.867$ ). The items are presented in [Textbox 2](#).

**Textbox 2.** Measurement of empathetic behavior.

<p>1: The physiotherapist behaved in such a way that the patient felt comfortable around her.</p> <p>2: The physiotherapist was caring and showed compassion to the patient.</p> <p>3: The physical therapist really listened to the patient.</p> <p>4: The physiotherapist encouraged the patient.</p> <p>5: The physiotherapist explained everything to the patient in an understandable way.</p>
---

We measured *attentiveness* with six self-created items. Three questions focused on the perceived attentiveness of the patient, and another three focused on the perceived attentiveness of the therapist ([Textbox 3](#)). The data were recorded using a 5-point

Likert scale. There was good internal consistency for both scales (attentiveness of the patient:  $\alpha=.830$ ; attentiveness of the physiotherapist:  $\alpha=.800$ ).

**Textbox 3.** Measurement of the perceived attentiveness of the physiotherapist and the patient. \*Reverse-coded items.

<p><b>Attentiveness of the physiotherapist</b></p> <ul style="list-style-type: none"> <li>-The therapist seemed to be distracted during the therapy.*</li> <li>-The therapist's attention was completely focused on the therapy.</li> <li>-The therapist showed great interest in the progress of the therapy.</li> </ul> <p><b>Attentiveness of the patient</b></p> <ul style="list-style-type: none"> <li>-The patient seemed to be distracted during the therapy.*</li> <li>-The patient's attention was completely focused on the therapy.</li> <li>-The patient showed great interest in the progress of the therapy.</li> </ul>
---

We also measured how participants evaluated the use of smartphones in physiotherapy in general (Textbox 4). The data were collected using a 5-point Likert scale. The internal consistency was good ( $\alpha=.895$ ).

**Textbox 4.** Evaluation of the usage of smartphones in physiotherapy. \*Reverse-coded items.

<p>1: A smartphone is mainly distracting in physiotherapy.*</p> <p>2: A smartphone interferes during physiotherapy.*</p> <p>3: A smartphone stands in the way of good physiotherapy.*</p> <p>4: It is very useful to use a smartphone in physiotherapy.</p> <p>5: The use of a smartphone in physiotherapy is very supportive.</p> <p>6: Using a smartphone in physiotherapy entails advantages.</p>
--

## Analysis

Data analysis was performed using IBM SPSS 25 statistics (IBM Corp) for Windows. To test our hypotheses and answer the open research questions, we performed contrast analysis. Contrast analysis allows for testing specific hypothesized patterns of mean differences by defining lambda coefficients while increasing statistical power and avoiding issues of multiple testing, which would arise with pairwise comparisons using *t* tests [65]. In addition, contrast analysis produces distinct test statistics for situations in which the group variances are equal or unequal. To decide for the correct test statistic, we used the Levene test for homogeneity of variance. To test the comparability of the conditions, ANOVA and chi-square tests were performed.

We provide mean, SD, and *R* as an indicator of effect size for significant results. According to a previous report [66], we interpreted  $R=0.10$  as a small effect,  $R=0.20$  as a typical effect, and  $R=0.30$  as a relatively large effect. The significance level for all analyses was set to  $\alpha=.05$ .

## Results

### Comparability of the Conditions

There were no significant differences among the four experimental conditions regarding participants' sex ( $\chi^2_3=2.02$ ,  $P=.57$ ), age ( $F_{3,114}=0.68$ ,  $P=.56$ ), and smartphone usage (average time of smartphone usage:  $P=.73$ ; time interval between getting up in the morning and the first use of the smartphone:  $P=.13$ ).

## Hypothesis Testing

All of the outcome variables (professional competence:  $P=.14$ ; social competence:  $P=.13$ ; attentiveness of the physiotherapist:  $P=.08$ ; attentiveness of the patient:  $P=.39$ ; smartphone use:  $P=.11$ ), except empathetic behavior ( $P=.01$ ), met the criteria of variance homogeneity. The scores for all of the experimental conditions are shown in Figure 3 and Figure 4.

In hypothesis 1, we assumed that the professional competence of the physiotherapist would be perceived as more pronounced if the physiotherapist used the smartphone in a purposeful way than if the physiotherapist looked at it from time to time with no therapeutic purpose. A contrast analysis supported this assumption ( $t_{112}=2.53$ ,  $P=.007$ ,  $R=0.23$ ). We found the same pattern of results for the perception of social competence (hypothesis 2;  $t_{112}=1.99$ ,  $P=.03$ ,  $R=0.19$ ) and empathy (hypothesis 3;  $t_{49,17}=1.82$ ,  $P=.04$ ,  $R=0.25$ ).

Hypothesis 4a was supported by the data as well. When the physiotherapist used the smartphone in a recreational way, the physiotherapist was perceived as being less attentive than in all of the other conditions ( $t_{105}=5.15$ ,  $P<.001$ ,  $R=0.45$ ). The same applied to hypothesis 4b that was concerned with the attentiveness of the patient. When the patient used the smartphone in a recreational way, the patient was also perceived as being less attentive than in all of the other conditions ( $t_{105}=7.63$ ,  $P<.001$ ,  $R=0.60$ ).

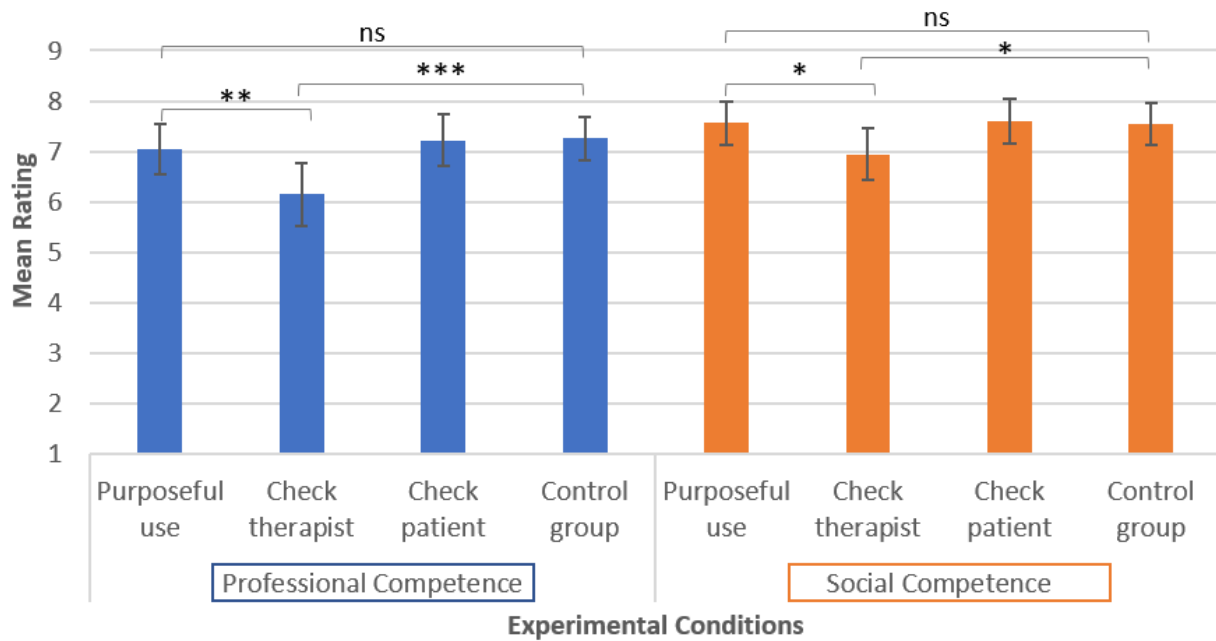
In line with hypothesis 5, we found that the way the smartphone was used had an impact on participants' evaluation of smartphone use in physiotherapy in general. If the physiotherapist used the smartphone in a purposeful way,



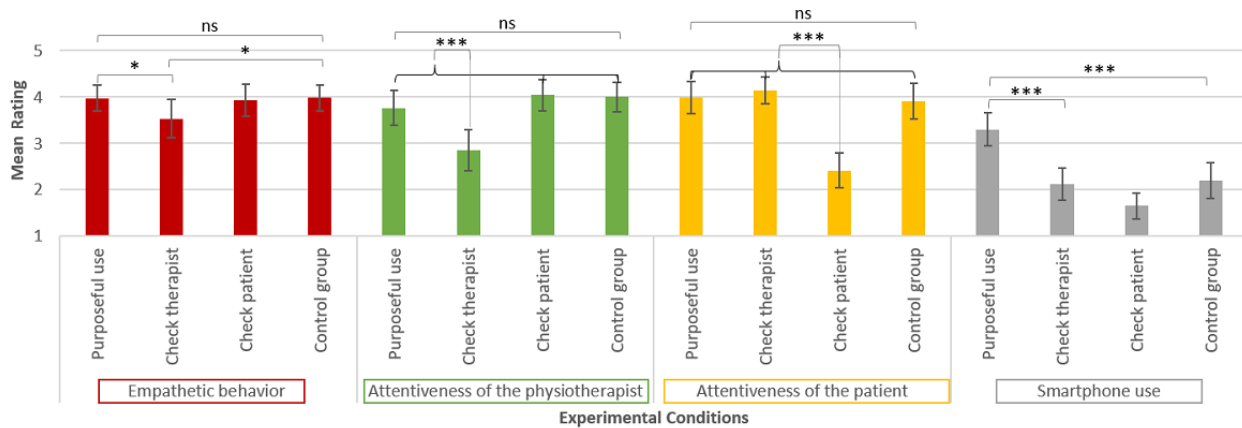
smartphone use in physiotherapy was rated more positively compared with the conditions in which the physiotherapist or

the patient looked at the smartphone from time to time with no therapeutic purpose ( $t_{102}=7.01, P<.001, R=0.57$ ).

**Figure 3.** Means and CIs for the outcome variables professional competence and social competence in the four experimental conditions. ns: nonsignificant. \* $P<.05$ , \*\* $P<.01$ , \*\*\* $P<.001$ .



**Figure 4.** Means and CIs for the outcome variables empathetic behavior, attentiveness of the physiotherapist, attentiveness of the patient, and smartphone use in the four experimental conditions. ns: nonsignificant. \* $P<.05$ , \*\*\* $P<.001$ .



**Open Research Questions**

As further contrast analyses showed, differences in perceived professional competence, perceived social competence, and perceived empathetic behavior occurred owing to the fact that the checking behavior of the physiotherapist led to a lower rating than no smartphone use (professional competence:  $t_{112}=3.12, P=.001, R=0.28$ ; social competence:  $t_{112}=1.93, P=.03, R=0.18$ ; empathy:  $t_{49.54}=1.86, P=.04, R=0.26$ ).

Purposeful use of the smartphone did not lead to a higher rating of perceived professional competence ( $P=.27$ ), perceived social competence ( $P=.48$ ), or perceived empathetic behavior ( $P=.47$ ) in comparison with no smartphone use. The positive evaluation of smartphone use in physiotherapy occurred owing to the fact

that purposeful use of the smartphone led to a more positive rating in comparison with no smartphone use ( $t_{102}=4.68, P<.001, R=0.42$ ) and not because use of the smartphone with no therapeutic purpose led to a less positive rating than no smartphone use ( $P=.38$ ).

**Discussion**

**Overview**

This study examined to what extent smartphone use in a physiotherapeutic treatment session impacted the perceived competence and empathy of the physiotherapist, the perceived attentiveness of the physiotherapist and the patient, and the overall evaluation of the use of smartphones in physiotherapy.

To our knowledge, this is the first empirical investigation into the effects of different kinds of smartphone usage in physiotherapeutic treatment situations. By examining the difference between purposeful use of smartphones and the often unconsciously occurring checking (or phubbing) behavior, we aimed to contribute to the highly socially relevant discussion about new interaction phenomena induced through the omnipresence of smartphones.

### General Findings

As expected, our results indicated that purposeful use of a smartphone by a physiotherapist was assessed differently from recreational smartphone use regarding the rating of the physiotherapist's professional and social competence and empathetic behavior. Interestingly, this was the case because the checking behavior of the physiotherapist led to a lower rating than no smartphone use and not because of the superiority of purposeful smartphone use. It seems that the smartphone used as a treatment tool was accepted, but it did not lead to any more positive perceptions of the therapist. It is therefore particularly important to use a smartphone in therapy only in a targeted manner; otherwise, it can have a negative impact on the physiotherapist-patient relationship, which is important for treatment success [46,54]. Regarding the perceived attentiveness of the physiotherapist and patient, the smartphone checking behavior again had a negative influence. Nevertheless, it was shown that a positive evaluation of smartphone usage in physiotherapy in general occurred owing to the fact that purposeful use of the smartphone led to a more positive rating in comparison with no smartphone use or checking behavior. This clearly showed that the participants saw the potential of using smartphones to improve treatment quality, but they also felt that inconsiderate use entailed hindrance of the social interaction in physiotherapy.

Overall, the results of our study are in line with other research findings demonstrating that smartphone use in social situations can have negative effects on social interactions [24,51,67] and perceived interpersonal attention [68]. However, the study also showed a possible remedy for the negative consequences of smartphone use, that is, utilizing the smartphone with a clear purpose can be experienced and appreciated by communication partners and patients.

### Limitations

Our research has some limitations worth noting. First, the experiment conducted here only relied on video material of a physiotherapeutic treatment situation. While videos can be a great format for illustrating processes, they are only an imitation of real treatment situations. Owing to the exact scripting of the video, we had the advantage of creating standardized experimental material, but the disadvantage of a reduced level of realism. In addition, the participants acted as observers and not as patients or therapists. We cannot know conclusively whether actual interaction partners in situations like the one in the video would perceive the situation in the same way. Moreover, this study cannot make a statement about whether a certain level of smartphone use is necessary to achieve certain effects, whether there must be a minimum threshold of smartphone use, or whether more smartphone use leads to stronger effects. The findings of this study only allow conclusions to be drawn about whether a certain type of use in general caused the effects investigated. Another limitation was the relatively small sample size. As the power analysis showed, we could only determine rather large effects. Large effects are also more relevant to clinical practice. Owing to the sample size, we cannot draw any conclusions on specific participant characteristics, such as gender and age. This should be addressed in further studies. It would also be interesting to examine different characteristics (eg, gender, age, and digital competence) of the therapist in further studies. Finally, we considered only one physiotherapeutic treatment situation. Generalization to other treatments or medical conditions is therefore not possible. We advise researchers to examine the effects of different kinds of smartphone uses in real clinical settings in the future.

### Conclusions

This research showed that the frequently reported negative effects of smartphone use in social (face-to-face) interactions were not perceived as such by our participants if the smartphone was used in an obviously purposeful way. In this kind of usage, participants even see the positive potential of smartphone use in physiotherapy. In summary, we recommend that practitioners use smartphones in a medical treatment situation only as actively integrated supportive tools for the purpose of treatment. Otherwise, a smartphone should be used with great awareness, and checking behavior should be avoided as often as possible.

---

### Acknowledgments

The research reported here was supported by budget resources of the Leibniz-Institut fuer Wissensmedien (Tuebingen, Germany).

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

CONSORT checklist.

[[PDF File \(Adobe PDF File\), 251 KB - mhealth\\_v9i4e25717\\_app1.pdf](#)]

---

### References

1. Rat C, Hild S, Rault Sérandour J, Gaultier A, Quereux G, Dreno B, et al. Use of Smartphones for Early Detection of Melanoma: Systematic Review. *J Med Internet Res* 2018 Apr 13;20(4):e135 [FREE Full text] [doi: [10.2196/jmir.9392](https://doi.org/10.2196/jmir.9392)] [Medline: [29653918](https://pubmed.ncbi.nlm.nih.gov/29653918/)]
2. Freeman K, Dinnes J, Chuchu N, Takwoingi Y, Bayliss SE, Matin RN, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. *BMJ* 2020 Feb 10;368:m127 [FREE Full text] [doi: [10.1136/bmj.m127](https://doi.org/10.1136/bmj.m127)] [Medline: [32041693](https://pubmed.ncbi.nlm.nih.gov/32041693/)]
3. Watson HA, Tribe RM, Shennan AH. The role of medical smartphone apps in clinical decision-support: A literature review. *Artif Intell Med* 2019 Sep;100:101707. [doi: [10.1016/j.artmed.2019.101707](https://doi.org/10.1016/j.artmed.2019.101707)] [Medline: [31607347](https://pubmed.ncbi.nlm.nih.gov/31607347/)]
4. Bergquist R, Vereijken B, Mellone S, Corzani M, Helbostad JL, Taraldsen K. App-based Self-administrable Clinical Tests of Physical Function: Development and Usability Study. *JMIR Mhealth Uhealth* 2020 Apr 27;8(4):e16507 [FREE Full text] [doi: [10.2196/16507](https://doi.org/10.2196/16507)] [Medline: [32338616](https://pubmed.ncbi.nlm.nih.gov/32338616/)]
5. Xu H, Long H. The Effect of Smartphone App-Based Interventions for Patients With Hypertension: Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth* 2020 Oct 19;8(10):e21759 [FREE Full text] [doi: [10.2196/21759](https://doi.org/10.2196/21759)] [Medline: [33074161](https://pubmed.ncbi.nlm.nih.gov/33074161/)]
6. Stubberud A, Linde M. Digital Technology and Mobile Health in Behavioral Migraine Therapy: a Narrative Review. *Curr Pain Headache Rep* 2018 Jul 31;22(10):66. [doi: [10.1007/s11916-018-0718-0](https://doi.org/10.1007/s11916-018-0718-0)] [Medline: [30066141](https://pubmed.ncbi.nlm.nih.gov/30066141/)]
7. Rathbone AL, Clarry L, Prescott J. Assessing the Efficacy of Mobile Health Apps Using the Basic Principles of Cognitive Behavioral Therapy: Systematic Review. *J Med Internet Res* 2017 Nov 28;19(11):e399 [FREE Full text] [doi: [10.2196/jmir.8598](https://doi.org/10.2196/jmir.8598)] [Medline: [29187342](https://pubmed.ncbi.nlm.nih.gov/29187342/)]
8. Kim Y, Wang Y, Oh J. Digital Media Use and Social Engagement: How Social Media and Smartphone Use Influence Social Activities of College Students. *Cyberpsychol Behav Soc Netw* 2016 Apr;19(4):264-269. [doi: [10.1089/cyber.2015.0408](https://doi.org/10.1089/cyber.2015.0408)] [Medline: [26991638](https://pubmed.ncbi.nlm.nih.gov/26991638/)]
9. Montag C, Blaszkiewicz K, Sariyska R, Lachmann B, Andone I, Trendafilov B, et al. Smartphone usage in the 21st century: who is active on WhatsApp? *BMC Res Notes* 2015 Aug 04;8:331 [FREE Full text] [doi: [10.1186/s13104-015-1280-z](https://doi.org/10.1186/s13104-015-1280-z)] [Medline: [26238512](https://pubmed.ncbi.nlm.nih.gov/26238512/)]
10. Kim JH, Park H. Effects of Smartphone-Based Mobile Learning in Nursing Education: A Systematic Review and Meta-analysis. *Asian Nurs Res (Korean Soc Nurs Sci)* 2019 Feb;13(1):20-29 [FREE Full text] [doi: [10.1016/j.anr.2019.01.005](https://doi.org/10.1016/j.anr.2019.01.005)] [Medline: [30659927](https://pubmed.ncbi.nlm.nih.gov/30659927/)]
11. Maudsley G, Taylor D, Allam O, Garner J, Calinici T, Linkman K. A Best Evidence Medical Education (BEME) systematic review of: What works best for health professions students using mobile (hand-held) devices for educational support on clinical placements? BEME Guide No. 52. *Med Teach* 2019 Feb;41(2):125-140. [doi: [10.1080/0142159X.2018.1508829](https://doi.org/10.1080/0142159X.2018.1508829)] [Medline: [30484351](https://pubmed.ncbi.nlm.nih.gov/30484351/)]
12. Kacetyl J, Klímová B. Use of Smartphone Applications in English Language Learning—A Challenge for Foreign Language Education. *Education Sciences* 2019 Jul 11;9(3):179. [doi: [10.3390/educsci9030179](https://doi.org/10.3390/educsci9030179)]
13. Dora J, van Hooff MLM, Geurts SAE, Hooftman WE, Kompier MAJ. Characterizing Work-Related Smartphone Use at Home and Private Smartphone Use at Work Using Latent Class Analysis. *Occup Health Sci* 2019 May 24;3(2):187-203. [doi: [10.1007/s41542-019-00040-6](https://doi.org/10.1007/s41542-019-00040-6)]
14. Park J, Kim S, Lee H. Effect of work-related smartphone use after work on job burnout: Moderating effect of social support and organizational politics. *Computers in Human Behavior* 2020 Apr;105:106194. [doi: [10.1016/j.chb.2019.106194](https://doi.org/10.1016/j.chb.2019.106194)]
15. Kushlev K, Dunn EW. Smartphones distract parents from cultivating feelings of connection when spending time with their children. *Journal of Social and Personal Relationships* 2018 Apr 10;36(6):1619-1639. [doi: [10.1177/0265407518769387](https://doi.org/10.1177/0265407518769387)]
16. Osborne R, Horberry T, Young KL. Pedestrian distraction from Smartphones: An end-user perspective on current and future countermeasures. *Transportation Research Part F: Traffic Psychology and Behaviour* 2020 Aug;73:348-361. [doi: [10.1016/j.trf.2020.07.007](https://doi.org/10.1016/j.trf.2020.07.007)]
17. Schneider FM, Hitzfeld S. I ought to put down that phone but I phub nevertheless: Examining the predictors of phubbing behavior. *Social Science Computer Review* 2019 Oct 24;089443931988236. [doi: [10.1177/0894439319882365](https://doi.org/10.1177/0894439319882365)]
18. Vorderer P, Hefner D, Reinecke L, Klimmt C. A New Paradigm in Communication Research? In: *Permanently Online, Permanently Connected*. Abingdon, United Kingdom: Routledge; 2017.
19. Vorderer P, Krömer N, Schneider FM. Permanently online – Permanently connected: Explorations into university students' use of social media and mobile smart devices. *Computers in Human Behavior* 2016 Oct;63:694-703. [doi: [10.1016/j.chb.2016.05.085](https://doi.org/10.1016/j.chb.2016.05.085)]
20. Antunovic D, Parsons P, Cooke TR. 'Checking' and googling: Stages of news consumption among young adults. *Journalism* 2016 Aug 08;19(5):632-648. [doi: [10.1177/1464884916663625](https://doi.org/10.1177/1464884916663625)]
21. Bayer JB, Campbell SW, Ling R. Connection Cues: Activating the Norms and Habits of Social Connectedness. *Commun Theor* 2015 Nov 24;26(2):128-149. [doi: [10.1111/comt.12090](https://doi.org/10.1111/comt.12090)]
22. Wilmer HH, Chein JM. Mobile technology habits: patterns of association among device usage, intertemporal preference, impulse control, and reward sensitivity. *Psychon Bull Rev* 2016 Oct;23(5):1607-1614. [doi: [10.3758/s13423-016-1011-z](https://doi.org/10.3758/s13423-016-1011-z)] [Medline: [26980462](https://pubmed.ncbi.nlm.nih.gov/26980462/)]

23. Chotpitayasunondh V, Douglas KM. How “phubbing” becomes the norm: The antecedents and consequences of snubbing via smartphone. *Computers in Human Behavior* 2016 Oct;63:9-18. [doi: [10.1016/j.chb.2016.05.018](https://doi.org/10.1016/j.chb.2016.05.018)]
24. Chotpitayasunondh V, Douglas KM. The effects of “phubbing” on social interaction. *J Appl Soc Psychol* 2018 Mar 25;48(6):304-316. [doi: [10.1111/jasp.12506](https://doi.org/10.1111/jasp.12506)]
25. Vanden Abeele MM, Hendrickson AT, Pollmann MM, Ling R. Phubbing behavior in conversations and its relation to perceived conversation intimacy and distraction: An exploratory observation study. *Computers in Human Behavior* 2019 Nov;100:35-47. [doi: [10.1016/j.chb.2019.06.004](https://doi.org/10.1016/j.chb.2019.06.004)]
26. Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. *Pers Ubiquit Comput* 2011 Jun 16;16(1):105-114. [doi: [10.1007/s00779-011-0412-2](https://doi.org/10.1007/s00779-011-0412-2)]
27. Duke É, Montag C. Smartphone addiction, daily interruptions and self-reported productivity. *Addict Behav Rep* 2017 Dec;6:90-95 [FREE Full text] [doi: [10.1016/j.abrep.2017.07.002](https://doi.org/10.1016/j.abrep.2017.07.002)] [Medline: [29450241](https://pubmed.ncbi.nlm.nih.gov/29450241/)]
28. Busch PA, McCarthy S. Antecedents and consequences of problematic smartphone use: A systematic literature review of an emerging research area. *Computers in Human Behavior* 2021 Jan;114:106414. [doi: [10.1016/j.chb.2020.106414](https://doi.org/10.1016/j.chb.2020.106414)]
29. Wang J, Wang P, Yang X, Zhang G, Wang X, Zhao F, et al. Fear of Missing Out and Procrastination as Mediators Between Sensation Seeking and Adolescent Smartphone Addiction. *Int J Ment Health Addiction* 2019 Jun 26;17(4):1049-1062. [doi: [10.1007/s11469-019-00106-0](https://doi.org/10.1007/s11469-019-00106-0)]
30. Horvath J, Mundinger C, Schmitgen MM, Wolf ND, Sambataro F, Hirjak D, et al. Structural and functional correlates of smartphone addiction. *Addict Behav* 2020 Jun;105:106334. [doi: [10.1016/j.addbeh.2020.106334](https://doi.org/10.1016/j.addbeh.2020.106334)] [Medline: [32062336](https://pubmed.ncbi.nlm.nih.gov/32062336/)]
31. Hassett L, van den Berg M, Weber H, Chagpar S, Wong S, Rabie A, et al. Activity and MObility UsiNg Technology (AMOUNT) rehabilitation trial - description of device use and physiotherapy support in the post-hospital phase. *Disabil Rehabil* 2020 Jul 14:1-7. [doi: [10.1080/09638288.2020.1790679](https://doi.org/10.1080/09638288.2020.1790679)] [Medline: [32663066](https://pubmed.ncbi.nlm.nih.gov/32663066/)]
32. Valdes K, Gendernalik E, Hauser J, Tipton M. Use of mobile applications in hand therapy. *J Hand Ther* 2020;33(2):229-234. [doi: [10.1016/j.jht.2019.10.003](https://doi.org/10.1016/j.jht.2019.10.003)] [Medline: [32253058](https://pubmed.ncbi.nlm.nih.gov/32253058/)]
33. Moral-Munoz JA, Zhang W, Cobo MJ, Herrera-Viedma E, Kaber DB. Smartphone-based systems for physical rehabilitation applications: A systematic review. *Assist Technol* 2019 May 21:1-14. [doi: [10.1080/10400435.2019.1611676](https://doi.org/10.1080/10400435.2019.1611676)] [Medline: [31112461](https://pubmed.ncbi.nlm.nih.gov/31112461/)]
34. Carroll WR, Bandura A. The role of visual monitoring in observational learning of action patterns: making the unobservable observable. *J Mot Behav* 1982 Jun;14(2):153-167. [doi: [10.1080/00222895.1982.10735270](https://doi.org/10.1080/00222895.1982.10735270)] [Medline: [15155177](https://pubmed.ncbi.nlm.nih.gov/15155177/)]
35. Heyes CM, Foster CL. Motor learning by observation: evidence from a serial reaction time task. *Q J Exp Psychol A* 2002 Apr;55(2):593-607. [doi: [10.1080/02724980143000389](https://doi.org/10.1080/02724980143000389)] [Medline: [12047061](https://pubmed.ncbi.nlm.nih.gov/12047061/)]
36. Rohbanfard H, Proteau L. Learning through observation: a combination of expert and novice models favors learning. *Exp Brain Res* 2011 Dec;215(3-4):183-197. [doi: [10.1007/s00221-011-2882-x](https://doi.org/10.1007/s00221-011-2882-x)] [Medline: [21986667](https://pubmed.ncbi.nlm.nih.gov/21986667/)]
37. Andrieux M, Proteau L. Observation learning of a motor task: who and when? *Exp Brain Res* 2013 Aug;229(1):125-137. [doi: [10.1007/s00221-013-3598-x](https://doi.org/10.1007/s00221-013-3598-x)] [Medline: [23748693](https://pubmed.ncbi.nlm.nih.gov/23748693/)]
38. Reo JA, Mercer VS. Effects of live, videotaped, or written instruction on learning an upper-extremity exercise program. *Phys Ther* 2004 Jul;84(7):622-633. [Medline: [15225081](https://pubmed.ncbi.nlm.nih.gov/15225081/)]
39. Miller J, Stanley I, Moore K. Videotaped exercise instruction: A randomised controlled trial in musculoskeletal physiotherapy. *Physiotherapy Theory and Practice* 2009 Jul 10;20(3):145-154. [doi: [10.1080/09593980490487375](https://doi.org/10.1080/09593980490487375)]
40. Rohbanfard H, Proteau L. Live vs. video presentation techniques in the observational learning of motor skills. *Trends in Neuroscience and Education* 2013 Mar;2(1):27-32. [doi: [10.1016/j.tine.2012.11.001](https://doi.org/10.1016/j.tine.2012.11.001)]
41. Post PG, Aiken CA, Laughlin DD, Fairbrother JT. Self-control over combined video feedback and modeling facilitates motor learning. *Hum Mov Sci* 2016 Jun;47:49-59. [doi: [10.1016/j.humov.2016.01.014](https://doi.org/10.1016/j.humov.2016.01.014)] [Medline: [26874750](https://pubmed.ncbi.nlm.nih.gov/26874750/)]
42. Coulson SE, Adams RD, O'Dwyer NJ, Crosson GR. Use of video self-modelling and implementation intentions following facial nerve paralysis. *International Journal of Therapy and Rehabilitation* 2006 Jan;13(1):30-35. [doi: [10.12968/ijtr.2006.13.1.21349](https://doi.org/10.12968/ijtr.2006.13.1.21349)]
43. Dowrick PW. A review of self modeling and related interventions. *Applied and Preventive Psychology* 1999 Dec;8(1):23-39. [doi: [10.1016/S0962-1849\(99\)80009-2](https://doi.org/10.1016/S0962-1849(99)80009-2)]
44. Hagin V, Gonzales BR, Gros Lambert A. Effects of cognitive stimulation with a self-modeling video on time to exhaustion while running at maximal aerobic velocity: a pilot study. *Percept Mot Skills* 2015 Apr;120(2):491-501. [doi: [10.2466/26.25.PMS.120v18x5](https://doi.org/10.2466/26.25.PMS.120v18x5)] [Medline: [25730197](https://pubmed.ncbi.nlm.nih.gov/25730197/)]
45. Steel KA, Adams RD, Coulson SE, Canning CG, Hawtin HJ. Video Self-model Training of Punt Kicking. *Int. J. Sport Health Sci* 2013;11:49-53. [doi: [10.5432/ijshs.201233](https://doi.org/10.5432/ijshs.201233)]
46. Bientzle M, Minje J, Cress U, Kimmerle J. Therapeutic Touch in Exercise Videos: A Randomized Experiment of the Impact on the Evaluation of Therapists' Competence and Viewers' Self-Reliance. *Front Sports Act Living* 2019 Sep 25;1:35 [FREE Full text] [doi: [10.3389/fspor.2019.00035](https://doi.org/10.3389/fspor.2019.00035)] [Medline: [33344958](https://pubmed.ncbi.nlm.nih.gov/33344958/)]
47. Bientzle M, Fissler T, Cress U, Kimmerle J. The impact of physicians' communication styles on evaluation of physicians and information processing: A randomized study with simulated video consultations on contraception with an intrauterine device. *Health Expect* 2017 Oct;20(5):845-851 [FREE Full text] [doi: [10.1111/hex.12521](https://doi.org/10.1111/hex.12521)] [Medline: [27860037](https://pubmed.ncbi.nlm.nih.gov/27860037/)]

48. Fissler T, Bientzle M, Cress U, Kimmerle J. The Impact of Advice Seekers' Need Salience and Doctors' Communication Style on Attitude and Decision Making: A Web-Based Mammography Consultation Role Play. *JMIR Cancer* 2015 Sep 08;1(2):e10 [FREE Full text] [doi: [10.2196/cancer.4279](https://doi.org/10.2196/cancer.4279)] [Medline: [28410160](https://pubmed.ncbi.nlm.nih.gov/28410160/)]
49. Ergün N, Göksu İ, Sakız H. Effects of Phubbing: Relationships With Psychodemographic Variables. *Psychol Rep* 2020 Oct;123(5):1578-1613. [doi: [10.1177/0033294119889581](https://doi.org/10.1177/0033294119889581)] [Medline: [31752605](https://pubmed.ncbi.nlm.nih.gov/31752605/)]
50. Al - Saggaf Y, O'Donnell SB. Phubbing: Perceptions, reasons behind, predictors, and impacts. *Human Behav and Emerg Tech* 2019 Apr 10;1(2):132-140. [doi: [10.1002/hbe2.137](https://doi.org/10.1002/hbe2.137)]
51. Nuñez TR, Radtke T, Eimler SC. A third-person perspective on phubbing: Observing smartphone-induced social exclusion generates negative affect, stress, and derogatory attitudes. *Cyberpsychology* 2020 Aug 10;14(3):3. [doi: [10.5817/cp2020-3-3](https://doi.org/10.5817/cp2020-3-3)]
52. Judd BK, Scanlan JN, Alison JA, Waters D, Gordon CJ. The validity of a professional competence tool for physiotherapy students in simulation-based clinical education: a Rasch analysis. *BMC Med Educ* 2016 Aug 05;16:196 [FREE Full text] [doi: [10.1186/s12909-016-0718-x](https://doi.org/10.1186/s12909-016-0718-x)] [Medline: [27492325](https://pubmed.ncbi.nlm.nih.gov/27492325/)]
53. Lo K, Curtis H, Keating JL, Bearman M. Physiotherapy clinical educators' perceptions of student fitness to practise. *BMC Med Educ* 2017 Jan 17;17(1):16 [FREE Full text] [doi: [10.1186/s12909-016-0847-2](https://doi.org/10.1186/s12909-016-0847-2)] [Medline: [28095836](https://pubmed.ncbi.nlm.nih.gov/28095836/)]
54. Kidd MO, Bond CH, Bell ML. Patients' perspectives of patient-centredness as important in musculoskeletal physiotherapy interactions: a qualitative study. *Physiotherapy* 2011 Jun;97(2):154-162. [doi: [10.1016/j.physio.2010.08.002](https://doi.org/10.1016/j.physio.2010.08.002)] [Medline: [21497250](https://pubmed.ncbi.nlm.nih.gov/21497250/)]
55. Allen MV, Roberts LC. Perceived acquisition, development and delivery of empathy in musculoskeletal physiotherapy encounters. *Journal of Communication in Healthcare* 2017 Aug 18;10(4):304-312. [doi: [10.1080/17538068.2017.1366000](https://doi.org/10.1080/17538068.2017.1366000)]
56. Rodríguez-Nogueira Ó, Moreno-Poyato AR, Álvarez-Álvarez MJ, Pinto-Carral A. Significant socio-emotional learning and improvement of empathy in physiotherapy students through service learning methodology: A mixed methods research. *Nurse Educ Today* 2020 Apr 19;90:104437. [doi: [10.1016/j.nedt.2020.104437](https://doi.org/10.1016/j.nedt.2020.104437)] [Medline: [32353642](https://pubmed.ncbi.nlm.nih.gov/32353642/)]
57. Ward A, Mandrusiak A, Levett-Jones T. Cultural empathy in physiotherapy students: a pre-test post-test study utilising virtual simulation. *Physiotherapy* 2018 Dec;104(4):453-461. [doi: [10.1016/j.physio.2018.07.011](https://doi.org/10.1016/j.physio.2018.07.011)] [Medline: [30318129](https://pubmed.ncbi.nlm.nih.gov/30318129/)]
58. Salo E, Salmela V, Salmi J, Numminen J, Alho K. Brain activity associated with selective attention, divided attention and distraction. *Brain Res* 2017 Jun 01;1664:25-36. [doi: [10.1016/j.brainres.2017.03.021](https://doi.org/10.1016/j.brainres.2017.03.021)] [Medline: [28363436](https://pubmed.ncbi.nlm.nih.gov/28363436/)]
59. Oviedo-Trespalacios O, Haque MM, King M, Washington S. Understanding the impacts of mobile phone distraction on driving performance: A systematic review. *Transportation Research Part C: Emerging Technologies* 2016 Nov;72:360-380. [doi: [10.1016/j.trc.2016.10.006](https://doi.org/10.1016/j.trc.2016.10.006)]
60. Yechiam E, Hochman G. Loss attention in a dual-task setting. *Psychol Sci* 2014 Feb;25(2):494-502. [doi: [10.1177/0956797613510725](https://doi.org/10.1177/0956797613510725)] [Medline: [24357614](https://pubmed.ncbi.nlm.nih.gov/24357614/)]
61. Boag RJ, Strickland L, Loft S, Heathcote A. Strategic attention and decision control support prospective memory in a complex dual-task environment. *Cognition* 2019 Oct;191:103974. [doi: [10.1016/j.cognition.2019.05.011](https://doi.org/10.1016/j.cognition.2019.05.011)] [Medline: [31234118](https://pubmed.ncbi.nlm.nih.gov/31234118/)]
62. Kok LM, Vlieland TPMV, Fiocco M, Nelissen RGHH. A comparative study on the prevalence of musculoskeletal complaints among musicians and non-musicians. *BMC Musculoskelet Disord* 2013 Jan 04;14:9 [FREE Full text] [doi: [10.1186/1471-2474-14-9](https://doi.org/10.1186/1471-2474-14-9)] [Medline: [23289849](https://pubmed.ncbi.nlm.nih.gov/23289849/)]
63. Willson P, McNamara JR. How perceptions of a simulated physician-patient interaction influence intended satisfaction and compliance. *Soc Sci Med* 1982;16(19):1699-1704. [doi: [10.1016/0277-9536\(82\)90095-8](https://doi.org/10.1016/0277-9536(82)90095-8)] [Medline: [7178916](https://pubmed.ncbi.nlm.nih.gov/7178916/)]
64. Neumann M, Wirtz M, Bollschweiler E, Warm M, Wolf J, Pfaff H. [Psychometric evaluation of the German version of the "Consultation and Relational Empathy" (CARE) measure at the example of cancer patients]. *Psychother Psychosom Med Psychol* 2008 Jan;58(1):5-15. [doi: [10.1055/s-2007-970791](https://doi.org/10.1055/s-2007-970791)] [Medline: [17429761](https://pubmed.ncbi.nlm.nih.gov/17429761/)]
65. Wiens S, Nilsson ME. Performing Contrast Analysis in Factorial Designs: From NHST to Confidence Intervals and Beyond. *Educ Psychol Meas* 2017 Aug;77(4):690-715 [FREE Full text] [doi: [10.1177/0013164416668950](https://doi.org/10.1177/0013164416668950)] [Medline: [29805179](https://pubmed.ncbi.nlm.nih.gov/29805179/)]
66. Gignac GE, Szodorai ET. Effect size guidelines for individual differences researchers. *Personality and Individual Differences* 2016 Nov;102:74-78. [doi: [10.1016/j.paid.2016.06.069](https://doi.org/10.1016/j.paid.2016.06.069)]
67. Nazir T, Pişkin M. Phubbing: A Technological Invasion Which Connected the World But Disconnected Humans. *International Journal of Indian Psychology* 2016;3(4):175-182. [doi: [10.25215/0304.195](https://doi.org/10.25215/0304.195)]
68. Przybylski AK, Weinstein N. Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. *Journal of Social and Personal Relationships* 2012 Jul 19;30(3):237-246. [doi: [10.1177/0265407512453827](https://doi.org/10.1177/0265407512453827)]

---

## Abbreviations

**ANOVA:** analysis of variance

---

*Edited by L Buis; submitted 22.11.20; peer-reviewed by J Redfern, H Xu, K Lo, DS Ko; comments to author 16.01.21; revised version received 01.02.21; accepted 10.03.21; published 21.04.21.*

*Please cite as:*

*Bientzle M, Restle A, Kimmerle J*

*Perception of Purposeful and Recreational Smartphone Use in Physiotherapy: Randomized Controlled Trial*

*JMIR Mhealth Uhealth 2021;9(4):e25717*

URL: <https://mhealth.jmir.org/2021/4/e25717>

doi: [10.2196/25717](https://doi.org/10.2196/25717)

PMID: [33881402](https://pubmed.ncbi.nlm.nih.gov/33881402/)

©Martina Bientzle, Anne Restle, Joachim Kimmerle. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 21.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Assessing the Quality of Mobile Health-Related Apps: Interrater Reliability Study of Two Guides

Jordi Miró<sup>1\*</sup>, PhD; Pere Llorens-Vernet<sup>1\*</sup>, PhD

Universitat Rovira i Virgili; Department of Psychology, Centre de Recerca en Avaluació i Mesura de la Conducta, Institut d'Investigació Sanitària Pere Virgili, Tarragona, Spain

\* all authors contributed equally

**Corresponding Author:**

Jordi Miró, PhD

Universitat Rovira i Virgili; Department of Psychology  
Centre de Recerca en Avaluació i Mesura de la Conducta  
Institut d'Investigació Sanitària Pere Virgili  
Dept de Psicologia, Carretera de Valls, s/n  
Tarragona, 43007  
Spain

Phone: 34 977558179

Email: [jordi.miro@urv.cat](mailto:jordi.miro@urv.cat)

## Abstract

**Background:** There is a huge number of health-related apps available, and the numbers are growing fast. However, many of them have been developed without any kind of quality control. In an attempt to contribute to the development of high-quality apps and enable existing apps to be assessed, several guides have been developed.

**Objective:** The main aim of this study was to study the interrater reliability of a new guide — the Mobile App Development and Assessment Guide (MAG) — and compare it with one of the most used guides in the field, the Mobile App Rating Scale (MARS). Moreover, we also focused on whether the interrater reliability of the measures is consistent across multiple types of apps and stakeholders.

**Methods:** In order to study the interrater reliability of the MAG and MARS, we evaluated the 4 most downloaded health apps for chronic health conditions in the medical category of IOS and Android devices (ie, App Store and Google Play). A group of 8 reviewers, representative of individuals that would be most knowledgeable and interested in the use and development of health-related apps and including different types of stakeholders such as clinical researchers, engineers, health care professionals, and end users as potential patients, independently evaluated the quality of the apps using the MAG and MARS. We calculated the Krippendorff alpha for every category in the 2 guides, for each type of reviewer and every app, separately and combined, to study the interrater reliability.

**Results:** Only a few categories of the MAG and MARS demonstrated a high interrater reliability. Although the MAG was found to be superior, there was considerable variation in the scores between the different types of reviewers. The categories with the highest interrater reliability in MAG were “Security” ( $\alpha=0.78$ ) and “Privacy” ( $\alpha=0.73$ ). In addition, 2 other categories, “Usability” and “Safety,” were very close to compliance (health care professionals:  $\alpha=0.62$  and  $0.61$ , respectively). The total interrater reliability of the MAG (ie, for all categories) was 0.45, whereas the total interrater reliability of the MARS was 0.29.

**Conclusions:** This study shows that some categories of MAG have significant interrater reliability. Importantly, the data show that the MAG scores are better than the ones provided by the MARS, which is the most commonly used guide in the area. However, there is great variability in the responses, which seems to be associated with subjective interpretation by the reviewers.

(*JMIR Mhealth Uhealth* 2021;9(4):e26471) doi:[10.2196/26471](https://doi.org/10.2196/26471)

**KEYWORDS**

mHealth; mobile health; mobile apps; evaluation studies, rating; interrater reliability; MARS; MAG

## Introduction

In recent years, there has been an explosion of interest in the use of mobile devices (eg, smartphones, tablets) [1], alongside huge advances in the development of health-related mobile apps [2]. For example, a total of 325,000 different health-related apps has recently been reported to be available [3]. There are mobile apps for virtually all kinds of health conditions: for example, chronic pain [4,5], cancer [6], diabetes [7], and cardiovascular diseases [8]. This growth has brought considerable benefits not only to patients but also to society at large and at multiple levels. For example, health-related apps help to (1) improve treatment management, (2) facilitate patient-doctor communication, (3) monitor the patient's condition in real time, and (4) improve accessibility to treatment [9-12]. But there is also a number of caveats, mostly related to the somewhat unsupervised and unregulated nature of the process. And it has been suggested that the fact that the field is evolving without much scientific support or guidance [13] not only acts as a barrier to improvement [14] but also, and more importantly, can potentially put an individual's health at risk [15]. Some of the main problems related to health apps are (1) faulty reminders that make proper treatment follow-up difficult (eg, the instructions on when to do an activity or take medication are not correct [16]); (2) lack of health expert involvement [17]; (3) inappropriate response to consumer needs (eg, bipolar disorder apps failing to provide any response when asked about extreme mood swings or suicidal ideation [18]); and (4) incorrect medication doses (eg, incorrect calculation of insulin dose from blood glucose values [19]).

In order to overcome the issues health-related apps are facing, some rating scales and guides have been developed (eg, [20,21]). One of the first was the Mobile App Rating Scale (MARS) [22]. It is one of the most used rating scales to measure the quality of health-related apps [23-27]. However, the MARS was created from a narrow perspective [28-30] on the basis of analyzing studies on existing mobile apps and leaving out information from other relevant sources (eg, standards governing the design of software for medical devices).

Recently, the Mobile App Development and Assessment Guide (MAG) [13] was created to address the problems observed in the guides available (but not current key concerns such as privacy and security) and to help assess health-related apps and guide stakeholders in the development of new quality apps. The MAG was developed using data from all potential relevant sources and a representative sample of the guidelines, frameworks, and standards in the field of health app development. The MAG has been acknowledged as a good quality guide by an international and interdisciplinary group of stakeholders [31].

These guides are important in the field as they provide quality scores that are key to identifying the best apps available and distinguishing them from the poorly designed ones. However, there are little data on the comparative value and consistency of the very few guides there are. The field would benefit considerably from studies that guide the development of new apps and comparatively assess the quality of existing ones.

The main objective of this research was to study and compare the MAG and MARS. More specifically, we aimed to compare the interrater reliability of the 2 measures. We also focused on whether the interrater reliability of the measures is consistent across multiple types of apps and stakeholders.

## Methods

### App Selection Process

In order to evaluate the interrater reliability of the MAG and MARS across different types of apps, we evaluated the top 4 search results for chronic health conditions in the medical category of the Apple and Android stores (ie, App Store and Google Play, respectively). The search and selection of the apps were conducted in October 2020.

The inclusion criteria were as follows: The app had to be focused on a chronic health condition, in English or Spanish, and free to download. We selected chronic health conditions because it is one of the domains in which health apps are becoming more relevant (56% of health apps are intended for this kind of patient [32]). Reports by governmental agencies indicate that chronic health conditions are a major health problem that affects 31% of the population [33-36]. In addition, chronic health conditions are the leading cause of death and disability in both the developed and developing world in the global burden of disease equation. The most important chronic health conditions are low back pain and headache, neoplasms, diabetes and kidney diseases, and cardiovascular diseases [37-40]. We used the following search terms, which are related to the top 4 chronic health conditions in the Global Burden of Disease study [41]: "pain," "cancer," "diabetes," and "cardiovascular." In this search, we identified 886 apps and excluded 265 as they were not related to any of the 4 health conditions of interest. Finally, we selected the top 4 most downloaded apps (1 for each chronic health condition), which we then used in this study.

### App Evaluation Process

The apps were rated by 8 reviewers during the months of October and November 2020. The reviewers were a group of stakeholders that included clinical researchers, engineers, health care professionals, and end users as potential patients. These groups of stakeholders were identified as representative of individuals that would be most knowledgeable and interested in the use and development of health-related apps. The individuals in the "end users/potential patients" and "health care professionals" groups were identified and approached by the authors while at the university hospital (for a health checkup or while at work, respectively). The individuals in the "clinical researchers" and "engineers" groups were professors or technicians working at the university. Only individuals that agreed to participate and reported having experience in the use of smartphones and health apps were selected. All individuals approached were included. Reviewers received (1) the list of apps, (2) a survey including the items of the MAG and MARS to be evaluated, and (3) specific instructions as to how to proceed with the review and evaluation of the apps. In order to avoid potential interferences and help reviewers to work independently, and in line with similar studies (eg, [42]), they



were not given any other suggestions, indications, or training about the procedure.

For the evaluation, all reviewers downloaded and installed the apps on their personal mobile device. Then, they reviewed each of the apps using the specific criteria in the MAG and MARS. In their assessment, the reviewers were instructed to only take into account the content and information provided within the app itself and the stores (ie, App Store and Google Play). This included websites, scientific studies, and other external references as long as they were suggested or mentioned explicitly within the app or the stores. Like similar successful procedures, the reviewers did not receive any specific training, and although they spent several minutes examining the apps, they were not instructed to use them realistically [42]. The objective of this activity and procedure was that they would evaluate the apps in the same way as experts who do not need them would.

The MAG [31] has 48 items grouped into 8 categories or domains: usability, privacy, security, appropriateness and suitability, transparency and content, safety, technical support and updates, and technology. The reviewers used each of the items in the categories to assess the quality of the apps and checked if the apps met those characteristics and functions (1=yes; 0=no).

The MARS [22] has 23 items that are grouped into 5 categories: engagement, functionality, aesthetics, information quality, and subjective quality. It also has 6 items that are app-specific and can be adapted to include or exclude specific information on the topic of interest. For example, these items have been used to assess the perceived effects on the user's knowledge, attitudes, and intentions to change as well as the likelihood of changing the identified targeted behaviors in a study of mobile apps supporting heart failure symptom monitoring and self-care management [23]. In this study, we discarded these app-specific items. When using the MARS, the reviewers used each of the items to assess the quality of the apps and scored them using a 5-point rating scale (1=inadequate, 2=poor, 3=acceptable, 4=good, 5=excellent).

## Data Analysis

In order to study and compare the interrater reliability of the MAG and MARS, we calculated the Krippendorff alpha [43,44] for every category in the 2 guides, for each kind of reviewer and every app, separately and combined. The Krippendorff coefficient has been found to be superior to the Cohen coefficient and can be used with an unlimited number of reviewers [45-47]. An alpha 0.667 has been identified as showing acceptable agreement [44]. Therefore, in this study, we used this figure as the minimum level showing agreement [44]. A negative alpha indicated that agreement was less than could be expected by chance. All data analyses were performed using SPSS v.26 for Windows using the Kalpha macro [48].

## Results

A total of 8 reviewers rated the 4 apps using the MAG and MARS guides. The mobile apps included in the analysis were "Manage My Pain" (ie, pain), "BELONG Beating Cancer Together" (ie, cancer), "mySugr - Diabetes App & Blood Sugar Tracker" (ie, diabetes), and "ASCVD Risk Estimator Plus" (ie, cardiovascular diseases).

The group of reviewers included 2 clinical researchers, 2 engineers, 2 health care professionals, and 2 end users as potential patients. Reviewers' ages ranged from 24 to 40 years old, with an equal distribution of women and men. Clinical researchers, engineers, and health care professionals had been involved in the development of health-related apps, but not in any of the apps and guides used in this study (they did not have any conflicts of interest). All reviewers were highly educated individuals (all had completed university studies) and were experienced smartphone and mobile app users.

Complete responses were provided for almost all criteria and apps, although a small number of criteria showed a percentage of data completeness that ranged from 78% to 97% (eg, "It has password management mechanisms"; see [Multimedia Appendix 1](#)). [Tables 1](#) and [2](#) show the interrater reliability coefficients by categories and overall for both guides.

**Table 1.** Interrater reliability scores when reviewers used the Mobile App Development and Assessment Guide (MAG).

Category	Reviewers				Aggregate
	Clinical researchers	Engineers	Health care professionals	End users	
Usability	0.28	0.28	0.62	0.45	0.38
Privacy	0.36	0.73	0.42	0.43	0.45
Security	0.18	0.78	0.76	0.26	0.47
Appropriateness and suitability	0.38	0	-0.15	0	0.25
Transparency and content	0	1	-0.40	-0.36	0.15
Safety	0.59	0.51	0.61	-0.23	0.33
Technical support and updates	0.38	1	1	0.76	0.30
Technology	0.44	0.45	-0.05	0.45	0.39
Total	0.40	0.66	0.55	0.29	0.45

**Table 2.** Interrater reliability scores when reviewers used the Mobile App Rating Scale (MARS).

Category	Reviewers				
	Clinical researchers	Engineers	Health care professionals	End users	Aggregate
Engagement	0.18	0.50	0.53	0.41	0.43
Functionality	0.24	0.52	0.40	-0.38	0.19
Aesthetics	0.42	0.26	0.23	-0.14	0.17
Information	0.03	0.08	0.05	-0.09	0.06
Subjective	0.57	0.41	-0.08	0.54	0.43
Total	0.27	0.41	0.25	0.19	0.29

For the MAG, the reviewers' scores for several categories complied with the criteria. The highest interrater reliability scores were for the categories "Privacy" (engineers:  $P=.73$ ) and "Security" (engineers:  $P=.78$ ; health care professionals:  $P=.76$ ). In addition, 2 other categories, "Usability" and "Safety," were very close to compliance (health care professionals:  $P=.62$  and  $P=.61$ , respectively). The total interrater reliability of MAG (ie, for all categories) was 0.45 (see [Table 1](#)).

For the MARS, none of the reviewers' scores or the aggregate scores complied with the criteria. The categories with the highest interrater index were "Engagement" and "Subjective" with an overall alpha coefficient of 0.43 in both cases. The total

interrater reliability of the MARS (ie, for all categories) was 0.29 (see [Table 2](#)).

[Tables 3](#) and [4](#) show the interrater reliability scores for each mobile app assessed using the MAG and MARS guides. As can be seen, none of the scores complied with the criteria overall or in any category. Nevertheless, the highest interrater reliability scores were for the MAG guide.

A comparison of the interrater reliability between MAG and MARS is shown in [Table 5](#). Additional supplementary information is also provided on the interrater reliability scores for each item (see [Multimedia Appendix 1](#)).

**Table 3.** Interrater reliability scores for apps when reviewers used the Mobile App Development and Assessment Guide (MAG).

Category	Mobile apps			
	Manage My Pain	BELONG Beating Cancer Together	mySugr - Diabetes App & Blood Sugar Tracker	ASCVD Risk Estimator Plus
Usability	0.58	0.49	0.27	0.15
Privacy	0.47	0.38	0.28	0.20
Security	0.44	0.18	0.42	0.32
Appropriateness and suitability	1	0.42	0	-0.04
Transparency and content	0.08	-0.08	-0.06	0.00
Safety	0	0.47	0.33	0.21
Technical support and updates	0.10	0.57	0.16	0.10
Technology	0.17	0.36	0.12	0.45
Total	0.53	0.42	0.32	0.35

**Table 4.** Interrater reliability scores for apps when reviewers used the Mobile App Rating Scale (MARS).

Category	Mobile apps			
	Manage My Pain	BELONG Beating Cancer Together	mySugr - Diabetes App & Blood Sugar Tracker	ASCVD Risk Estimator Plus
Engagement	0.31	0.24	-0.10	0.18
Functionality	0.27	0.05	-0.02	0.16
Aesthetics	-0.05	-0.03	-0.07	0.12
Information	-0.08	0.08	-0.03	0.09
Subjective	0.55	0.44	0.16	0.14
Total	0.20	0.18	0.01	0.42

**Table 5.** Interrater reliability scores of the Mobile App Development and Assessment Guide (MAG) and the Mobile App Rating Scale (MARS).

Guide and category	Reliability
<b>MAG</b>	
Usability	0.38
Privacy	0.45
Security	0.47
Appropriateness and suitability	0.25
Transparency and content	0.15
Safety	0.33
Technical support and updates	0.30
Technology	0.39
Total	0.45
<b>MARS</b>	
Engagement	0.43
Functionality	0.19
Aesthetics	0.17
Information	0.06
Subjective	0.43
Total	0.29

## Discussion

### Principal Findings

This research is the first to measure the interrater reliability of the MAG [13,31]. We used the MAG to study 4 mobile health-related apps and compared the results with those obtained with the MARS [22], one of the most extensively used guides in the field.

In studies using the Krippendorff alpha, it is customary to require an alpha >0.800. However, an alpha 0.667 has been identified as indicative of acceptable agreement, and anything below that is considered as unacceptable [42,44]. The data revealed that few categories reached that score and showed high interrater reliability. This finding is similar to that of other studies (eg, [26,42,46]) that have analyzed this type of guides. Taken as a whole, the findings demonstrate that it is difficult for reviewers to rate the apps in the same or similar way. First of all, reviewers greatly differed in the amount of time spent in reviewing each app (ranging from 30 minutes to 60 minutes). Thus, it is possible that the time spent in the review process had an influence on the results of the assessment. Our data did not show differences associated to the amount of time spent in the review. However, we used a small number of reviewers (n=8). Therefore, additional research to study this issue is warranted. Another potential explanation for the findings is that reviewers do not interact with the apps in the same way, so they display different responses and functions [46]. Therefore, it is unlikely that reviewers will detect all app functions, which leads to differences in the ratings because they might not be assessing exactly the same items. Support for this explanation can be found in the fact that the most objective categories evaluated,

those which require less subjective interpretation by the reviewer (eg, “Privacy,” “Security”), are the ones with the highest interrater reliabilities. This finding is similar to the one reported by Powell and colleagues [42], who detected that the less judgment required by reviewers, the higher the reliability.

Another important finding of this study was that interrater reliability scores for the MAG were better than for the MARS. Importantly, some of the MAG categories with the highest interrater reliability are not included in the MARS (eg, “Privacy,” “Security,” “Technology”). These are issues that have grown in importance in the field in recent years.

It should also be noted that some MAG categories showed a higher interrater reliability than others, but there was considerable variation in the scores between the types of reviewer. This finding suggests that the differences in the interrater reliability scores are related to such individual characteristics of the reviewers as background or training. This could help explain, in part at least, why engineers showed the highest reliability scores in the category of “Security,” as this is an important issue that is currently a matter of key interest in the training of engineers but not in the case of clinical researchers. And it implies that reviewers from different backgrounds are required to assess apps and that reviewers need to be trained. However, it is also possible that the low interrater reliability scores were not only reviewer-related, but also app-related. That is, although we selected the 4 most downloaded apps, they may not have been quality apps or easy to assess (eg, the functions or properties of the apps were not easy to find or identify). In support of this explanation, some items were not answered by any reviewers in either of the guides (eg, “It has a data recovery system in case of loss”; “It is based on ethical principles and values”). Finally, another nonexclusive

explanation for these results could be related to the guides (ie, the MAG and MARS). The fact that the categories that required less interpretation (eg, “Security”) were the ones with the highest interrater reliability would support this explanation. This suggests that the guides must be improved.

The differences in interrater reliability and more importantly, the lower scores found suggest that there is a very important underlying problem that is indicative of the difficulty of creating a good guide to help in the development and assessment of health-related apps. On the basis of the results of this study and others (eg, [42,46]), users of health-related apps should use and interpret the results of quality assessments with caution. The guides, as they are, have not been demonstrated to provide a secure reliable measure of the overall quality of the apps.

The assessment of the quality of health-related apps is very important. Therefore, we must continue working on improving the way assessments are conducted. This may not only require improving the available guides but also working with specialized centers and trained reviewers.

### Future Research

Studies are needed to help improve available guides that are psychometrically sound so future research should focus on how to improve and empirically test interrater reliability. For example, studies should examine whether giving reviewers additional training is enough or how reviewers’ knowledge and assessment skills can best be improved. They should also establish whether the quality of health-related apps should be assessed by reviewers with different qualifications, training, and background. Moreover, since subjectivity might be an issue in the guides, an area for improvement is that guides include clearly defined criteria. Therefore, research to determine whether understandable and well-defined criteria can improve interrater reliability above and beyond the improvement in reviewer training is warranted. Moreover, and specifically in relation with the MAG, additional research with more apps of different types is also warranted. This would help ascertain whether and how different types of app influence the reviewers’ evaluations. In addition, the criteria and the categories included in the guide deserve specific attention. Studies with additional samples of reviewers, including individuals with chronic health conditions, to evaluate their comprehensibility and appropriateness are needed.

### Acknowledgments

This work was partly supported by grants from the Spanish Ministry of Economy, Industry and Competitiveness (RTI2018-09870-B-I00; RED2018-102546-T); European Regional Development Fund, the Government of Catalonia (AGAUR; 2017SGR-1321); Fundació Grünenthal (Spain), Universitat Rovira i Virgili (PFR program); and ICREA-Acadèmia. PL benefitted from a predoctoral fellowship (2019 FI\_B2 00151) cofinanced by the Secretaria d’Universitats i Recerca del Departament d’Empresa i Coneixement de la Generalitat de Catalunya, the European Union, and the European Social Fund.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Interrater reliability scores and data completeness for each item.

### Limitations

This study has a number of limitations that should be taken into account when interpreting the results. First, we studied the interrater reliability of the MAG when it was used to evaluate apps that were available for both Android and IOS. Although the apps are generally the same on both platforms, there may be small differences that influence the user’s experience or performance when using different platforms and devices. For example, the amount of information displayed or the position and size of some elements (eg, buttons, menu) may differ due to the size of the screen. Second, we used a very limited number of apps. We selected the most downloaded ones, as we thought they would be of better quality and therefore easier for reviewers to assess. However, they may not be of quality or representative of health-related apps and so may not be suitable for an accurate study of the interrater reliability of the guides. Third, during the period of time that the apps were being assessed, they may have been updated or modified, which would have had an unknown impact on the results of the assessments. Fourth, although individuals from different groups participated, they may be not representative. Even though they were extremely knowledgeable in their respective areas, they may or may not be the best individuals to assess the quality of the apps, as none of them had received any training. Moreover, they did not receive any substantial training in using the MAG or MARS. Thus, it is unclear whether the low interrater reliability is related to the instrument that is being used, to the lack of training provided to the raters, or both. We decided not to give specific training as we wanted to study whether the MAG and MARS can be reliably used as they are. Previous studies have also used this strategy (eg, [42]). However, future studies should examine whether training can help improve the reviewers’ assessment and the interrater reliability.

### Conclusions

Despite the limitations of the study, our findings provide new and important information about the MAG. Of particular consequence is that several categories in the MAG have significant interrater reliability. In addition, the data show that the scores are better than the ones provided by the MARS, the most commonly used guide in the area.

[PDF File (Adobe PDF File), 142 KB - [mhealth\\_v9i4e26471\\_app1.pdf](#) ]

## References

1. The Mobile Economy 2019. GSMA Intelligence. 2019. URL: <https://www.gsmainelligence.com/research/?file=b9a6e6202ee1d5f787cfebb95d3639c5&download> [accessed 2020-07-14]
2. Statista. App stores: number of apps in leading app stores 2020. 2020. URL: <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/> [accessed 2020-07-14]
3. mHealth Economics 2017: Current Status and Future Trends in Mobile Health. Research2Guidance. 2017. URL: <https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health> [accessed 2020-07-14]
4. de la Vega R, Roset R, Castarlenas E, Sánchez-Rodríguez E, Solé E, Miró J. Development and testing of painometer: a smartphone app to assess pain intensity. *J Pain* 2014 Oct;15(10):1001-1007. [doi: [10.1016/j.jpain.2014.04.009](https://doi.org/10.1016/j.jpain.2014.04.009)] [Medline: [24854065](https://pubmed.ncbi.nlm.nih.gov/24854065/)]
5. de la Vega R, Roset R, Galán S, Miró J. Fibroline: A mobile app for improving the quality of life of young people with fibromyalgia. *J Health Psychol* 2018 Jan;23(1):67-78. [doi: [10.1177/1359105316650509](https://doi.org/10.1177/1359105316650509)] [Medline: [27225285](https://pubmed.ncbi.nlm.nih.gov/27225285/)]
6. Kessel KA, Vogel MM, Schmidt-Graf F, Combs SE. Mobile Apps in Oncology: A Survey on Health Care Professionals' Attitude Toward Telemedicine, mHealth, and Oncological Apps. *J Med Internet Res* 2016 Nov 24;18(11):e312 [FREE Full text] [doi: [10.2196/jmir.6399](https://doi.org/10.2196/jmir.6399)] [Medline: [27884810](https://pubmed.ncbi.nlm.nih.gov/27884810/)]
7. Kebede MM, Pischke CR. Popular Diabetes Apps and the Impact of Diabetes App Use on Self-Care Behaviour: A Survey Among the Digital Community of Persons With Diabetes on Social Media. *Front. Endocrinol* 2019 Mar 1;10:135 [FREE Full text] [doi: [10.3389/fendo.2019.00135](https://doi.org/10.3389/fendo.2019.00135)] [Medline: [30881349](https://pubmed.ncbi.nlm.nih.gov/30881349/)]
8. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. *Eur J Prev Cardiol* 2018 Mar;25(5):505-521. [doi: [10.1177/2047487317750913](https://doi.org/10.1177/2047487317750913)] [Medline: [29313363](https://pubmed.ncbi.nlm.nih.gov/29313363/)]
9. Global Observatory for eHealth series - Volume 3. World Health Organization. 2011 Jun 07. URL: [https://www.who.int/goe/publications/ehealth\\_series\\_vol3/en/](https://www.who.int/goe/publications/ehealth_series_vol3/en/) [accessed 2021-04-10]
10. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile Apps for Increasing Treatment Adherence: Systematic Review. *J Med Internet Res* 2019 Jun 18;21(6):e12505 [FREE Full text] [doi: [10.2196/12505](https://doi.org/10.2196/12505)] [Medline: [31215517](https://pubmed.ncbi.nlm.nih.gov/31215517/)]
11. Ernesting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmeier A, et al. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. *J Med Internet Res* 2017 Apr 05;19(4):e101 [FREE Full text] [doi: [10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)] [Medline: [28381394](https://pubmed.ncbi.nlm.nih.gov/28381394/)]
12. 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):e52 [FREE Full text] [doi: [10.2196/jmir.3951](https://doi.org/10.2196/jmir.3951)] [Medline: [25803266](https://pubmed.ncbi.nlm.nih.gov/25803266/)]
13. Llorens-Vernet P, Miró J. Standards for Mobile Health-Related Apps: Systematic Review and Development of a Guide. *JMIR Mhealth Uhealth* 2020 Mar 03;8(3):e13057 [FREE Full text] [doi: [10.2196/13057](https://doi.org/10.2196/13057)] [Medline: [32130169](https://pubmed.ncbi.nlm.nih.gov/32130169/)]
14. de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. a systematic review of pain-related apps. *PLoS One* 2014;9(7):e101312 [FREE Full text] [doi: [10.1371/journal.pone.0101312](https://doi.org/10.1371/journal.pone.0101312)] [Medline: [24999983](https://pubmed.ncbi.nlm.nih.gov/24999983/)]
15. 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](https://doi.org/10.1093/jamia/ocz175)] [Medline: [31599936](https://pubmed.ncbi.nlm.nih.gov/31599936/)]
16. Park JYE, Li J, Howren A, Tsao NW, De Vera M. Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews. *JMIR Mhealth Uhealth* 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: [10.2196/11919](https://doi.org/10.2196/11919)] [Medline: [30702435](https://pubmed.ncbi.nlm.nih.gov/30702435/)]
17. Subhi Y, Bube SH, Rolskov BS, Skou TAS, 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](https://doi.org/10.2196/mhealth.4169)] [Medline: [26215371](https://pubmed.ncbi.nlm.nih.gov/26215371/)]
18. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality. *J Med Internet Res* 2015;17(8):e198 [FREE Full text] [doi: [10.2196/jmir.4581](https://doi.org/10.2196/jmir.4581)] [Medline: [26283290](https://pubmed.ncbi.nlm.nih.gov/26283290/)]
19. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. *BMC Med* 2015;13:106 [FREE Full text] [doi: [10.1186/s12916-015-0314-7](https://doi.org/10.1186/s12916-015-0314-7)] [Medline: [25943590](https://pubmed.ncbi.nlm.nih.gov/25943590/)]
20. Medicines and Healthcare products Regulatory Agency. Medical devices: software applications (apps). GOV.UK. 2014 Aug 08. URL: <https://www.gov.uk/government/publications/medical-devices-software-applications-apps> [accessed 2020-07-14]
21. Good practice guidelines on health apps and smart devices (mobile health or mhealth). Haute Autorité de Santé. 2016. URL: [https://www.has-sante.fr/jcms/c\\_2681915/en/good-practice-guidelines-on-health-apps-and-smart-devices-mobile-health-or-mhealth](https://www.has-sante.fr/jcms/c_2681915/en/good-practice-guidelines-on-health-apps-and-smart-devices-mobile-health-or-mhealth) [accessed 2020-07-14]

22. 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](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
23. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS). *JMIR Mhealth Uhealth* 2016 Jun 14;4(2):e74 [FREE Full text] [doi: [10.2196/mhealth.5882](https://doi.org/10.2196/mhealth.5882)] [Medline: [27302310](https://pubmed.ncbi.nlm.nih.gov/27302310/)]
24. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone Apps Targeting Alcohol and Illicit Substance Use: Systematic Search in in Commercial App Stores and Critical Content Analysis. *JMIR Mhealth Uhealth* 2019 Apr 22;7(4):e11831 [FREE Full text] [doi: [10.2196/11831](https://doi.org/10.2196/11831)] [Medline: [31008713](https://pubmed.ncbi.nlm.nih.gov/31008713/)]
25. Mehdi M, Stach M, Riha C, Neff P, Dode A, Pryss R, et al. Smartphone and Mobile Health Apps for Tinnitus: Systematic Identification, Analysis, and Assessment. *JMIR Mhealth Uhealth* 2020 Aug 18;8(8):e21767 [FREE Full text] [doi: [10.2196/21767](https://doi.org/10.2196/21767)] [Medline: [32808939](https://pubmed.ncbi.nlm.nih.gov/32808939/)]
26. 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](https://doi.org/10.2196/14991)] [Medline: [31381501](https://pubmed.ncbi.nlm.nih.gov/31381501/)]
27. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for People With Rheumatoid Arthritis to Monitor Their Disease Activity: A Review of Apps for Best Practice and Quality. *JMIR Mhealth Uhealth* 2017 Feb 21;5(2):e7 [FREE Full text] [doi: [10.2196/mhealth.6956](https://doi.org/10.2196/mhealth.6956)] [Medline: [28223263](https://pubmed.ncbi.nlm.nih.gov/28223263/)]
28. Nouri R, R Niakan Kalhori S, 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](https://doi.org/10.1093/jamia/ocy050)] [Medline: [29788283](https://pubmed.ncbi.nlm.nih.gov/29788283/)]
29. Siddique AB, Krebs M, Alvarez S, Greenspan I, Patel A, Kinsolving J, et al. Mobile Apps for the Care Management of Chronic Kidney and End-Stage Renal Diseases: Systematic Search in App Stores and Evaluation. *JMIR Mhealth Uhealth* 2019 Sep 04;7(9):e12604 [FREE Full text] [doi: [10.2196/12604](https://doi.org/10.2196/12604)] [Medline: [31486408](https://pubmed.ncbi.nlm.nih.gov/31486408/)]
30. Levine DM, Co Z, Newmark LP, Groisser AR, Holmgren AJ, Haas JS, et al. Design and testing of a mobile health application rating tool. *NPJ Digit Med* 2020;3:74 [FREE Full text] [doi: [10.1038/s41746-020-0268-9](https://doi.org/10.1038/s41746-020-0268-9)] [Medline: [32509971](https://pubmed.ncbi.nlm.nih.gov/32509971/)]
31. 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](https://doi.org/10.2196/17760)] [Medline: [32735226](https://pubmed.ncbi.nlm.nih.gov/32735226/)]
32. mHealth Economics 2016 - Current Status and Trends of the mHealth App Market. Research2Guidance. 2016. URL: <https://research2guidance.com/product/mhealth-app-developer-economics-2016/> [accessed 2020-07-14]
33. Chronic Diseases in America. Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/chronicdisease/resources/infographic/chronic-diseases.htm> [accessed 2020-07-14]
34. Chronic Disease Overview. Australian Institute of Health and Welfare. 2019. URL: <https://www.aihw.gov.au/reports-data/health-conditions-disability-deaths/chronic-disease/overview> [accessed 2020-07-14]
35. EU research on chronic diseases. European Commission. 2014. URL: [https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/chronic-diseases\\_en](https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/chronic-diseases_en) [accessed 2020-07-14]
36. Health at a Glance 2019. Organisation for Economic Cooperation and Development (OECD). 2019. URL: [https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019\\_4dd50c09-en](https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019_4dd50c09-en) [accessed 2021-04-10]
37. Findings from the Global Burden of Disease Study 2017. Institute for Health Metrics and Evaluation (IHME). 2018. URL: [http://www.healthdata.org/sites/default/files/files/policy\\_report/2019/GBD\\_2017\\_Booklet.pdf](http://www.healthdata.org/sites/default/files/files/policy_report/2019/GBD_2017_Booklet.pdf) [accessed 2020-07-14]
38. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7)] [Medline: [30496103](https://pubmed.ncbi.nlm.nih.gov/30496103/)]
39. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Dec 10;392(10159):1789-1858 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)] [Medline: [30496104](https://pubmed.ncbi.nlm.nih.gov/30496104/)]
40. GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1859-1922 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32335-3](https://doi.org/10.1016/S0140-6736(18)32335-3)] [Medline: [30415748](https://pubmed.ncbi.nlm.nih.gov/30415748/)]
41. Global Burden of Disease Study 2017 (GBD 2017) Data Resources. Institute for Health Metrics and Evaluation (IHME). 2017. URL: <http://ghdx.healthdata.org/gbd-2017> [accessed 2020-07-14]
42. Powell AC, Torous J, Chan S, Raynor GS, Shwartz E, Shanahan M, et al. Interrater Reliability of mHealth App Rating Measures: Analysis of Top Depression and Smoking Cessation Apps. *JMIR Mhealth Uhealth* 2016 Feb 10;4(1):e15 [FREE Full text] [doi: [10.2196/mhealth.5176](https://doi.org/10.2196/mhealth.5176)] [Medline: [26863986](https://pubmed.ncbi.nlm.nih.gov/26863986/)]
43. Krippendorff K. Computing Krippendorff's Alpha-Reliability. University of Pennsylvania. 2011. URL: [https://repository.upenn.edu/asc\\_papers/43](https://repository.upenn.edu/asc_papers/43) [accessed 2020-07-14]

44. Krippendorff K. Content Analysis: An Introduction to Its Methodology. Thousand Oaks, CA: Sage Publications, Inc; 2018.
45. Karlsson D, Gøeg KR, Öрман H, Højen AR. Semantic Krippendorff's  $\alpha$  for measuring inter-rater agreement in SNOMED CT coding studies. Stud Health Technol Inform 2014;205:151-155. [Medline: [25160164](#)]
46. 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](#)]
47. Volkmann N, Stracke J, Kemper N. Evaluation of a gait scoring system for cattle by using cluster analysis and Krippendorff's  $\alpha$  reliability. Vet Rec 2019 Feb 16;184(7):220. [doi: [10.1136/vr.105059](#)] [Medline: [30455190](#)]
48. Hayes AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data. Communication Methods and Measures 2007;1(1):77-89. [doi: [10.1080/19312450709336664](#)]

## Abbreviations

**MAG:** Mobile App Development and Assessment Guide

**MARS:** Mobile App Rating Scale

*Edited by L Buis; submitted 13.12.20; peer-reviewed by W Schlee, L Guo; comments to author 27.01.21; revised version received 29.01.21; accepted 20.03.21; published 19.04.21.*

*Please cite as:*

Miró J, Llorens-Vernet P

Assessing the Quality of Mobile Health-Related Apps: Interrater Reliability Study of Two Guides

JMIR Mhealth Uhealth 2021;9(4):e26471

URL: <https://mhealth.jmir.org/2021/4/e26471>

doi: [10.2196/26471](#)

PMID: [33871376](#)

©Jordi Miró, Pere Llorens-Vernet. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 19.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Twelve-Lead Electrocardiogram Acquisition With a Patchy-Type Wireless Device in Ambulance Transport: Simulation-Based Randomized Controlled Trial

Sunyoung Yoon<sup>1</sup>, BSc; Taerim Kim<sup>2</sup>, MD, PhD; Taehwan Roh<sup>3</sup>, PhD; Hansol Chang<sup>1,2</sup>, MD; Sung Yeon Hwang<sup>2</sup>, MD, PhD; Hee Yoon<sup>2</sup>, MD, PhD; Tae Gun Shin<sup>2</sup>, MD, PhD; Min Seob Sim<sup>2</sup>, MD, PhD; Ik Joon Jo<sup>2</sup>, MD, PhD; Won Chul Cha<sup>1,2,4</sup>, MD

<sup>1</sup>Department of Digital Health, Samsung Advanced Institute for Health Science & Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

<sup>2</sup>Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

<sup>3</sup>Healthrian Co, Ltd, Dajeon, Republic of Korea

<sup>4</sup>Health Information and Strategy Center, Samsung Medical Center, Seoul, Republic of Korea

**Corresponding Author:**

Won Chul Cha, MD

Department of Emergency Medicine

Samsung Medical Center

Sungkyunkwan University School of Medicine

81 Irwon-ro, Gangnam-gu

Seoul, 06351

Republic of Korea

Phone: 82 2 3410 2053

Email: [docchaster@gmail.com](mailto:docchaster@gmail.com)

## Abstract

**Background:** Cardiovascular disease is the leading cause of death worldwide. Early recognition, diagnosis, and reperfusion are the key elements of treatment for ST-segment elevation myocardial infarction. The absence of a prehospital 12-lead electrocardiogram (P12ECG) can cause definitive treatment delay and repeated transfer. Although guidelines highly recommend the measurement and transmission of P12ECG data, P12ECG use has not been widely established.

**Objective:** The aim of this study was to verify the time-efficiency and feasibility of the use of a patchy-type 12-lead ECG measuring and transmitting device (P-ECG) by an emergency medical technician (EMT) in an ambulance during patient transport.

**Methods:** This was a simulation-based prospective randomized crossover-controlled study that included EMTs. The participants were randomly assigned to one of two groups. Group A began the experiment with a conventional 12-lead ECG (C-ECG) device and then switched to the intervention device (P-ECG), whereas group B began the experiment with the P-ECG and then switched to the C-ECG. All simulations were performed inside an ambulance driving at 30 km/h. The time interval was measured from the beginning of ECG application to completion of sending the results. After the simulation, participants were administered the System Usability Scale questionnaire about usability of the P-ECG.

**Results:** A total of 18 EMTs were recruited for this study with a median age of 35 years. The overall interval time for the C-ECG was 254 seconds (IQR 247-270), whereas the overall interval time for the P-ECG was 130 seconds (IQR 112-150), with a significant difference ( $P < .001$ ). Significant differences between the C-ECG and P-ECG were identified at all time intervals, in which the P-ECG device was significantly faster in all intervals, except for the preparation interval in which the C-ECG was faster ( $P = .03$ ).

**Conclusions:** Performance of 12-lead ECG examination and transmission of the results using P-ECG are faster than those of C-ECG during ambulance transport. With the additional time afforded, EMTs can provide more care to patients and transport patients more rapidly, which may help reduce the symptoms-to-balloon time for patients with acute coronary syndrome.

**Trial Registration:** ClinicalTrials.gov NCT04114760; <https://www.clinicaltrials.gov/ct2/show/NCT04114760>

(*JMIR Mhealth Uhealth* 2021;9(4):e24142) doi:[10.2196/24142](https://doi.org/10.2196/24142)



**KEYWORDS**

12-lead electrocardiogram; electrocardiogram transmission; prehospital; wearable patch device; wearable; electrocardiogram; ECG; cardiovascular; efficiency; feasibility; EMT

## Introduction

### Background and Significance

Cardiovascular disease (CVD) is the leading cause of death worldwide [1]. In 2017, there were 17.9 million deaths resulting from CVD, more than three-quarters of which occurred in low- or middle-income countries [2]. Deaths caused by ischemic heart disease have risen by about 19%-20% over the past 10 years. Although age-standardized death rates from CVD have shown a decreasing trend globally, CVD remains one of the primary causes of death [3]. To reduce the total ischemic time, as a major factor in short- and long-term mortality [1,4,5], early recognition, diagnosis, and reperfusion should be performed in a coordinated and complementary manner.

### Prehospital 12-Lead Electrocardiogram

The prehospital 12-lead electrocardiogram (P12ECG) is a cornerstone of emergency cardiac care. Capturing and transmitting P12ECG data to the emergency department is an integral part of patient care for acute coronary syndrome (ACS) [6], and serves as an inexpensive, noninvasive diagnostic tool. Both the American Heart Association and the European Society of Cardiology recommend the capture and transmission of P12ECG data for patients who present with symptoms suggestive of ACS [1,4]. Missing P12ECG in patients with suspected ACS is a risk factor for delay in reperfusion treatment or transfer to a nondesignated hospital [7].

However, use of P12ECG has not been widely established. A large retrospective cohort study performed in the United States found that before hospital arrival, only 32% of patients had received P12ECG, and among the patients who visited the emergency department via a 911 call with a final confirmed diagnosis of ACS, almost 41% had not received prehospital electrocardiogram (ECG) [8]. Other studies have also indicated that P12ECG data are transmitted to the hospital for only a small portion of patients. A study performed in Los Angeles found that among those diagnosed with ST-segment elevation myocardial infarction (STEMI) by P12ECG, only 28% had their data transmitted to the hospital; in 55% of cases, the data were not transmitted and for 17% of these cases the transmission status was not recorded [9]. In another study performed in Poland, 12-lead ECG was transmitted to the hospital by an emergency medical service (EMS) team for only 2% of patients [10].

Identifiable reasons for missing P12ECG include transmission malfunction, too short transfer interval for performing P12ECG, and only a 3-lead ECG available [11]. The most frequently used 12-lead ECG examination device during the prehospital stage and in the ambulance is the defibrillator; however, this device can transfer ECG data only by registered email [12]. Various handheld and wearable ECG devices [13,14] and transmission technologies [15-19] have been developed in recent years, but few studies have examined both the device application and

transmission of P12ECG in an ambulance during patient transport.

### Study Objective

The aim of this study was to verify the time-efficiency and feasibility of the use of a patchy-type 12-lead ECG measuring and transmitting device (P-ECG) by emergency medical technicians (EMTs) in an ambulance during patient transport.

## Methods

### Study Design

This was a simulation-based prospective randomized crossover-controlled study that included EMTs. The experiment was performed in an ambulance driving 30 km/h. The experiment was performed at a single institution and the driving course was limited to the hospital site. This study used a case-crossover design and participants were randomly assigned to one of two groups.

This study was approved by the Samsung Medical Center Institutional Review Board (No. 2019-04-004) and is registered at ClinicalTrials.gov (NCT04114760).

### Study Setting

The EMS system in Korea consists of 18 provincial headquarters, 224 fire stations, and 1474 ambulance stations. There are 12,033 EMTs registered in the system who are classified into level 1 or level 2 certification, which determines their scope. The scope of procedures and skills corresponding to level-1 EMTs in Korea are comparable to those of EMTs at the intermediate level in North America. Intermediate/level-1 EMTs can perform a 12-lead ECG examination with physician oversight.

### Study Population and Sample Size

The inclusion criteria were as follows: aged 19 years or older and a level-1 certified EMT. The exclusion criteria were as follows: not working as a field agent (eg, a dispatcher) and only a level-2 EMT license. Study recruitment posters were posted in a fire station in Seoul, and participants were recruited through voluntary participation. All participants were given a reward worth approximately US \$20.

The sample size was calculated based on our unpublished pilot study involving hospital health care providers. Participants were divided into two groups: the conventional 12-lead ECG (C-ECG) group and the P-ECG group. The mean time required in the C-ECG group was 161.35 seconds (SD 81.57) and the mean time required in the P-ECG group was 79.12 seconds (SD 24.75). The Wilcoxon signed-rank test was used to compare the means between paired samples with a crossover design. The minimum number of samples needed for hypothesis testing was calculated with a power of 0.95, effect size of 1.13, and type I error of .05. The effect size was estimated based on previous studies and calculated using G power 3.1.9.4. [20]. Based on

the above information, it was determined that the minimum number of participants needed for this study was 13. Assuming a 40% dropout rate, 18 was the target number of participants.

Data on the demographic and study-related characteristics of the participants were collected, including gender and age, previous hospital training, and number of years as an EMT. In addition, education and training were provided to the participants by a professor of emergency medicine. The training included instruction on proper 12-lead ECG methods, introduction to the P-ECG, practice, and a question and answer session.

**Materials**

The ZOLL X series defibrillator (Chelmsford, MA, USA) was used as the C-ECG device and wearECG12 (HEALTHRIAN, Daejeon, Republic of Korea) was used as the P-ECG device. The X-series is 22.6x26.4x20.1 cm and weighs 5.3 kg (including the battery and paper). The X-series is currently the most commonly used equipment in the Korean EMS system and is included among the first-aid equipment mounted in the ambulance. The device is capable of both defibrillation and ECG monitoring. A mode setting, cable change, and 10-electrode attachment are required for the 12-lead ECG examination, and the results are printed on paper (Figure 1).

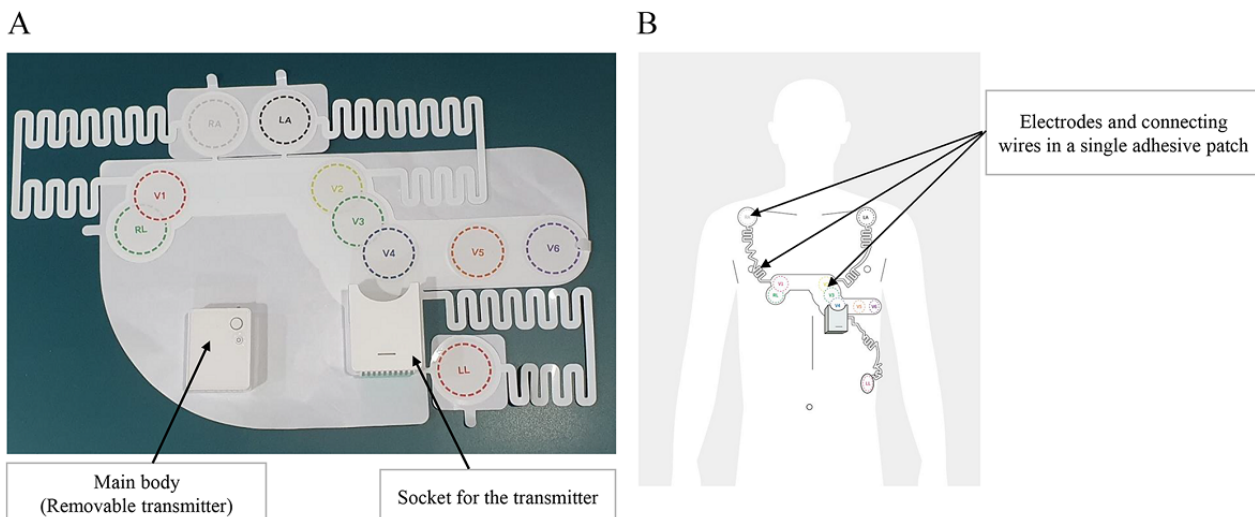
**Figure 1.** (A) The conventional 12-lead electrocardiogram device used for the control condition (X-series, ZOLL Medical, Chelmsford, MA, USA). (B) Attachment of electrodes on the patient’s body surface, which are connected to the main device.



The P-ECG, which was used as the intervention device, was approved by the Korean Ministry of Food and Drug Administration as Holter ECG technology (Figure 2). This device consists of a main body (46x35.6x16 mm, 30 g) and a one-patch type electrode (241.19x375.5 mm, 35 g). The main body is assembled on the socket of the patch. For performing 12-lead ECG examinations, the tablet and main body are wirelessly connected via Bluetooth (Kirkland, WA, USA) and provide continuous monitoring. The results of an ECG

examination are generated in a PDF that was transmitted in real time to the researchers’ dashboards via long-term evolution (LTE) networks when the tablet’s “upload” button was clicked. The tablet used in this study was Samsung Galaxy Tab S3 (SM-T825; Seoul, Republic of Korea) on an LTE network. Samsung Galaxy Books (SM-W627NZFKOO) was used as the dashboard on an LTE network. The differences between the two ECG devices are summarized in Table 1.

**Figure 2.** (A) Patchy-type wireless 12-lead electrocardiogram used for the intervention condition (HEALTHRIAN, Yuseong-gu, Daejeon, Republic of Korea). The device consists of two parts: a flexible patch with a socket and a transmitter. (B) Application of the device with the transmitter on the patient’s left chest area.



**Table 1.** Functional comparison between the conventional 12-lead electrocardiogram (C-ECG) and patchy-type wireless 12-lead electrocardiogram (P-ECG) devices.

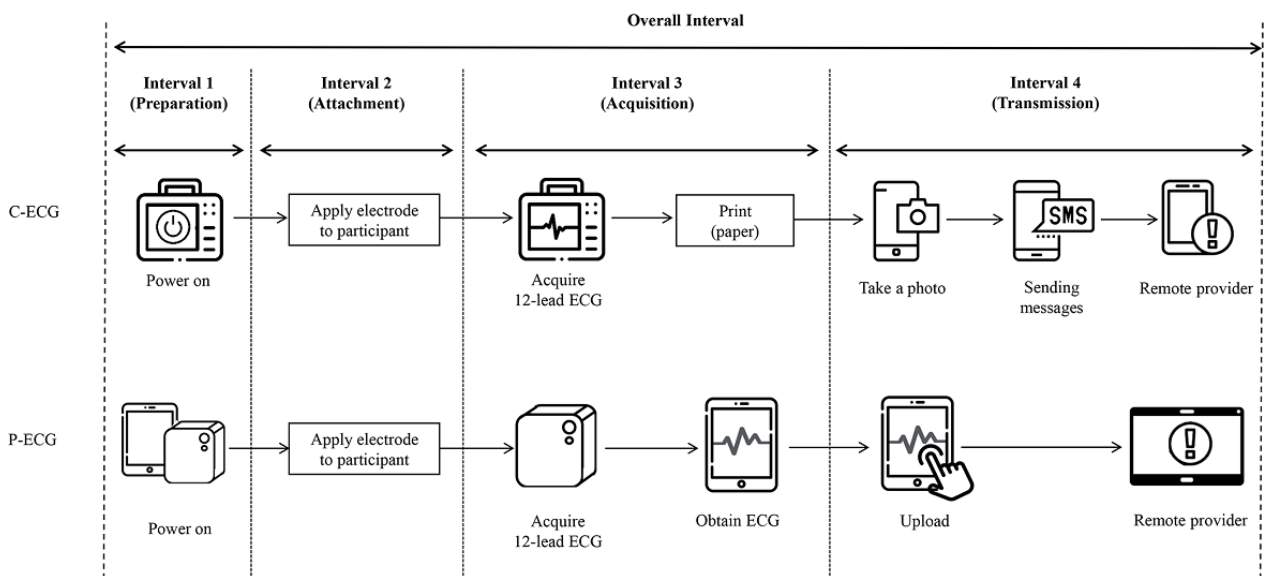
Component	C-ECG	P-ECG
Electrode	10 electrodes (separated)	Single electrode (combined as a patch)
Wire	10 wires for each electrode	Wireless (conductive film on the patch)
Control	Manual control (minimum 10 steps)	Semiautomatic (4 steps)
Transmission	Messenger app (device supports email with a cellular dongle)	Built-in app

### Study Protocol

All participants provided written informed consent before the start of the simulation. They were provided with 15 minutes of education on 12-lead ECG and the two devices, followed by 10 minutes of practice. All participants already knew how to use the C-ECG and therefore only practiced using the P-ECG. They were randomly assigned to one of two groups. Group A began the experiment with the C-ECG and then transferred to the P-ECG, whereas group B began the experiment with the P-ECG

and then transferred to the C-ECG. Both groups performed the 12-lead ECG examination on a simulated mock patient. The volunteers were selected according to the inclusion criteria of healthy men aged 19 years or older. Six healthy men aged 28 to 35 years were included as mock patients. Each mock patient received an ECG examination six times by the EMT participants. Testing was performed in the same order for both devices: preparation, attachment, acquisition, and transmission (Figure 3).

**Figure 3.** Definition of outcomes (intervals). The overall interval was defined as the interval from the "start" command to the acquisition of an electrocardiogram (ECG) image by a remote provider. C-ECG: conventional electrocardiogram; P-ECG: patchy-type electrocardiogram.



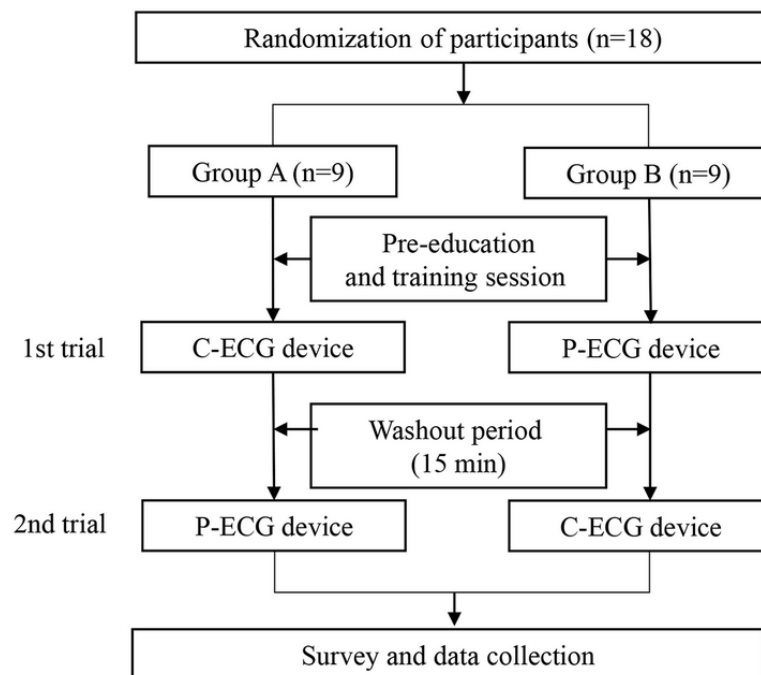
The following procedure was used for the C-ECG (Figure 3): (1) preparation, turn on the device and change the mode and cable to that for 12-lead ECG; (2) attachment, apply the 10 electrodes and the wire cable, and modify the positions of the electrodes as needed so that the four limb leads are located on both the upper shoulder and lower abdomen (Figure 1B); (3) acquisition, perform the 12-lead ECG examination and print the results; (4) transmission, take a photo of the printout of the results and send it in a message to the researcher's cell phone.

The procedure for the P-ECG was as follows (Figure 3): (1) preparation, turn on the main body of the device, assemble it

on the patch's socket, and complete Bluetooth pairing; (2) attachment, apply the one-patch type electrodes and modify the electrode positions as needed (Figure 2B); (3) acquisition, perform the 12-lead ECG examination and obtain the ECG results (PDF); (4) transmission, complete the transmission of the results to the researcher's dashboard using the LTE network by touching the "upload" button on the tablet.

After the simulations, a modified version of the System Usability Scale (SUS) [21] was administered to all participants to assess their satisfaction with the patchy-type wireless device (Figure 4).

**Figure 4.** Case-crossover design study process. There was a washout period before two trials. C-ECG: conventional 12-lead electrocardiogram; P-ECG: patchy-type wireless 12-lead electrocardiogram.



## Outcome Measures

The primary outcomes of the study were the time intervals during each stage (Figure 3). Interval 1 was the preparation time, from powering on to having the system ready for ECG acquisition. Interval 2 was the time for attachment, measured from the beginning of electrode attachment to final modification of the position of the attached electrodes. Interval 3 was the time taken to acquire a complete 12-lead ECG for transmission. Interval 4 was the time for transmission, from completion of 12-lead ECG results acquisition to completion of transmission to a distant health care provider. The overall interval was calculated as the sum of intervals 1 to 4.

All steps were compared and verified using a stopwatch and video recording to ensure time interval accuracy. The video recording did not reveal the faces of the participants, and as stated on the consent form, the video was only used for research purposes.

The secondary outcome of the study was the SUS score. This outcome was used to investigate participant satisfaction with the P-ECG device. In this study, the SUS was administered only for the P-ECG. There were two reasons for assessing the usability of only one device. First, this study was designed as a case-crossover study. All participants used both devices and the order of experience was randomized. Having mixed experiences could have made it difficult for participants to clearly respond to two surveys at the same time. Second, all participants were already familiar with the C-ECG, which is among the required first-aid equipment in the ambulance. Therefore, the score for the conventional device would likely

have been much higher regardless of the usability of the device itself. The SUS consists of 10 questions, each based on a 5-point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). The formula used to calculate the SUS score was as follows [21]:



The SUS score was evaluated based on a previous study [22], which interpreted mean SUS scores above 12.5 as “Worst Imaginable,” above 20.3 as “Awful,” above 35.7 as “Poor,” above 50.9 as “OK,” above 71.4 as “Good,” above 85.5 as “Excellent,” and above 90.9 as “Best imaginable.”

## Data Analysis

The differences in paired values of the time required for the same participant were analyzed using the Wilcoxon signed-rank test. *P* values less than .05 were considered statistically significant. All data processing and statistical analyses were performed using R version 3.6.3 software (R Foundation for Statistical Computing, Vienna, Austria).

## Results

### Participant Characteristics

Table 2 presents the demographic and study-related characteristics of the participants. The median age of the participants was 35 years. The median continuous years of service as an EMT was 8 years and 5 months (IQR 4 years, 7 months to 13 years, 1 month). About half of all participants had worked in a tertiary hospital for more than 2 years before serving in the EMT (Table 2).

**Table 2.** Demographic and study-related characteristics of the participants (N=18).

Characteristic	Value
<b>Gender, n (%)</b>	
Women	4 (22)
Men	14 (78)
Age (years), median (IQR)	35 (32-42)
<b>Previous training in the hospital, n (%)</b>	
Yes	10 (56)
No	8 (44)
<b>Years of service in the emergency medical service, n (%)</b>	
0-4	4 (22)
5-8	5 (28)
9-18	9 (50)

### Time Interval Comparison

The overall interval time for the C-ECG device was significantly slower than that for the P-ECG device (Table 3). Significant

differences between the C-ECG device and the P-ECG device were identified at all time intervals, in which the P-ECG device was significantly faster for all intervals except interval 1 (Table 3).

**Table 3.** Comparison of time intervals between the conventional electrocardiogram (C-ECG) and patchy-type electrocardiogram (P-ECG) devices.

Interval	C-ECG, median (IQR)	P-ECG, median (IQR)	<i>P</i> value <sup>a</sup>
Interval 1 (preparation)	26 (22-33)	34 (31-47)	.03
Interval 2 (attachment)	77 (64-91)	53 (43-75)	.03
Interval 3 (acquisition)	69 (66-75)	24 (21-30)	<.001
Interval 4 (transmission)	74 (58-97)	6 (4-8)	<.001
Overall interval	254 (247-270)	130 (112-150)	<.001

<sup>a</sup>*P* values were calculated using the Wilcoxon signed-rank test.

### SUS Survey Outcomes

According to a previous study, SUS scores higher than 71.4 can be interpreted as “good” [22]. In this study, the total SUS score

was 73.75, indicating that the P-ECG has good usability and is acceptable to the user (Table 4).

**Table 4.** System Usability Scale adapted for assessment of participant satisfaction with the patchy-type wireless electrocardiogram (ECG) device.

Question	Mean (SD) <sup>a</sup>
I think that I would like to use the ECG device frequently	3.33 (0.91)
I found the ECG device to be unnecessarily complex	2.50 (1.15)
I thought the ECG device was easy to use	3.33 (0.84)
I think that I would need the support of a technical person to be able to use the ECG device	2.39 (1.33)
I found that the various functions in the ECG device were well-integrated	2.78 (0.94)
I thought there was too much inconsistency in the ECG device	3.00 (0.69)
I would imagine that most people would learn to use the ECG device very quickly	3.50 (0.79)
I found the ECG device very cumbersome to use	3.22 (1.11)
I felt very confident using the ECG device	3.06 (0.64)
I need to learn a lot of things before I could get going with the ECG device	2.39 (1.14)
Total score	73.75 (17.58)

<sup>a</sup>Scored on a scale from 1 (“strongly disagree”) to 5 (“strongly agree”).

## Discussion

### Principal Results

To the best of our knowledge, our study is the first to investigate a new P12ECG device and system using an LTE network and a single patchy-type 12-lead ECG, transmitting real-time 12-lead ECG data to dispatcher dashboards during ambulance transport. Except for during the preparation stage, the intervention device was at least 20 seconds faster during all stages and was approximately 70 seconds faster during the transmission stage. As a result, the total time from preparation to transmission completion was reduced from 4 minutes to 2 minutes, despite this being the first use of an unfamiliar device.

Although a 2-minute difference may seem trivial, this corresponds to nearly half of the conventional mean. With that additional time, EMTs can provide much needed care to the patient and transport the patient more rapidly to the destination hospital. Although it was not determined in this study, the new system requires very little interaction once attached, and the 2 minutes of application time required for the new device may be reduced given the EMT's free hands.

A prior study found that the main reason for failure of performed or transmitted P12ECG was a short transfer time to the hospital [11], suggesting that reducing the total required time of P12ECG by half would improve the P12ECG measurement and transmission rate within 10 minutes of first medical contact. In addition, reducing 2 minutes in the prehospital field is more critical than in the in-hospital setting because delays in testing can result in delays in transport and definitive care.

### Downsides of the New Device

There has been a relative decline in ACS mortality trends in all national regions; however, the prevalence rates differ depending on the country's gross domestic product per capita [23]. ACS mortality trends are also influenced by various factors such as the patient's medical condition (eg, the presence of diabetes mellitus, renal failure, or previous coronary artery disease) and performance of the STEMI network based on the emergency medical and primary health care system [1]. Although P12ECG transmission is considered a cornerstone, the cost and necessary regional infrastructure to maintain such activity could be a burden in many regions.

Although not proven in this study, there could be two major challenges to the wide adoption of P12ECG. The first is the price. Conventional ECG uses very cheap, disposable electrodes, and the new device's patch includes electrodes and wire-like films, which are also disposable, making its use more expensive. The second is the placement of electrodes on the surface of the body; the new device's lib leads are placed in a more proximal area compared with conventional devices, which makes it difficult to compare ECG readings from different ECG types.

### Potential of the New Device

In the emergency medical system, particularly during the prehospital stage, an EMS-friendly device is required. ECG technology was developed about 120 years ago by Einthoven. The initial model required patients to place their extremities in

buckets for measurement. The present 12-lead ECG was invented about 80 years ago by Dr. Emanuel [24] and is only used inside of a hospital in early states. Although some studies have found the prehospital and in-hospital quality of 12-lead ECG to be equal [25-28], it is difficult to actively apply large and heavy defibrillator devices in the prehospital environment. Most potential ACS patients are alert and rarely need defibrillation immediately; therefore, the advantage of separating a 12-lead ECG device from the defibrillator is expected to increase the use of P12ECG.

Well-planned and well-coordinated programs that provide sufficient training are essential for the effective implementation of P12ECG [29]. During the 15-minute pretraining session used in our study, most participants took some time to get used to the new device and system. The lack of sufficient training time could increase the total time as the primary outcome of this study. The SUS results suggest that this might have been the case. The question with the highest rate of agreement was related to believing that most people would learn to use the ECG device very quickly, whereas the questions related to evaluating device function as being well-integrated and feeling confident using the ECG device had the highest rate of disagreement. That is, participants believed they could learn to use the device quickly but they were not yet familiar with the new device function and were not confident after the short training they received for the study. Despite the short training time and low confidence, the overall time was reduced by nearly half in this study, and it is highly likely that more familiarity with P-ECG, leading to equivalent confidence in using C-ECG, would increase the difference in the overall time interval, eventually making the outcome more convincing.

Current technology offers a more enhanced user experience and is more usable compared with previous devices. A novel, smart device shortens the learning period through its simpler and more intuitive design. Although the use of new devices and systems may initially cause discomfort and lower confidence, there may be sufficient improvement with training.

### Optimized Networking System for Hardware and Software

Easy and fast transmission methods are also essential in applying P12ECG. Because real-time decision-making is needed, EMTs share ECG data with a cell phone camera, messaging app, or via a verbal report. These methods require EMTs to have high ECG interpretation abilities, and the use of an unencrypted mobile device can compromise personal information. This leads to issues regarding Health Insurance Portability and Accountability Act data breaches [30]. A safe and easy way to transfer information from devices, mobile devices, or a connected app is needed.

For the treatment of CVD, an effective and efficient way to integrate prehospital care into in-hospital care should be sought to improve the overall quality of care. Simple 12-lead ECG measurement and easy transmission would not only contribute to the quality of prehospital emergency care but also improve hospital care. In the future, these technologies are expected to extend to patients' homes, enabling 12-lead ECG management of CVD by the patients themselves.

## Limitations

This study had some limitations. First, this was a simulation-based study, and the driving course and speed of the ambulance were limited to within the hospital area. The ambulance was only able to reach 30 km/h. Furthermore, our simulated patient was a healthy adult man. Second, the study location was in Seoul, where both LTE and 5G networks are well-implemented, and the average age of participants was 36 years. Therefore, transmission-related problems and bias may have occurred in other areas related to the use of mobile devices and app manipulation. Third, because only one device was used, it is difficult to represent the characteristics of other devices used in the prehospital environment. Environmental factors such as humidity can affect the touch sensitivity function of tablets and can cause recognition problems in the actual hospital environment. Fourth, this study collected 12-lead ECG data but

only assessed time intervals and SUS survey responses as outcomes, without a comparison of the quality and accuracy of the two different ECG devices. Finally, although many studies have reported that reduction of total ischemic time may improve short- and long-term mortality, and reduce complications such as heart failure [31-33], this study only examined one component of STEMI treatment. Therefore, further research on the emergency medical system is needed.

## Conclusions

The performance and transmission of the 12-lead ECG examination using a patchy-type 12-lead ECG is faster than those when using the conventional device during ambulance transport. With this additional time, EMTs can provide more care to patients and transport patients more rapidly, which may help to reduce the symptoms-to-balloon time for patients with ACS.

## Acknowledgments

This work was supported by an Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (2020-0-00224, Development of wireless multi-lead electrocardiogram monitoring device and cardiac diagnosis solution). We thank the EMTs from Gangnam Fire Station in Seoul for their enthusiastic participation in the study.

## Conflicts of Interest

TR is the CEO of HEALTHRIAN, which developed wearECG12 that was used as our intervention device in the study. The other authors have no conflicts of interest to declare.

## References

1. Ibanez B, James S, Agewall S, Antunes M, Bucciarelli-Ducci C, Bueno H, ESC Scientific Document Group. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J* 2018 Jan 07;39(2):119-177. [doi: [10.1093/eurheartj/ehx393](https://doi.org/10.1093/eurheartj/ehx393)] [Medline: [28886621](https://pubmed.ncbi.nlm.nih.gov/28886621/)]
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7)] [Medline: [30496103](https://pubmed.ncbi.nlm.nih.gov/30496103/)]
3. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017 Sep 16;390(10100):1151-1210 [FREE Full text] [doi: [10.1016/S0140-6736\(17\)32152-9](https://doi.org/10.1016/S0140-6736(17)32152-9)] [Medline: [28919116](https://pubmed.ncbi.nlm.nih.gov/28919116/)]
4. O'Gara P, Kushner F, Ascheim D, Casey DJ, Chung M, de Lemos JA, American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation* 2013 Jan 29;127(4):e362-e425. [doi: [10.1161/CIR.0b013e3182742cf6](https://doi.org/10.1161/CIR.0b013e3182742cf6)] [Medline: [23247304](https://pubmed.ncbi.nlm.nih.gov/23247304/)]
5. Studencan M, Alusik D, Plachy L, Bajerovska L, Ilavsky M, Karas J, et al. Significant benefits of new communication technology for time delay management in STEMI patients. *PLoS One* 2018;13(11):e0205832 [FREE Full text] [doi: [10.1371/journal.pone.0205832](https://doi.org/10.1371/journal.pone.0205832)] [Medline: [30388116](https://pubmed.ncbi.nlm.nih.gov/30388116/)]
6. Zègre-Hemsey JK, Drew BJ. Prehospital electrocardiography: a review of the literature. *J Emerg Nurs* 2012 Jan;38(1):9-14 [FREE Full text] [doi: [10.1016/j.jen.2011.09.001](https://doi.org/10.1016/j.jen.2011.09.001)] [Medline: [22137883](https://pubmed.ncbi.nlm.nih.gov/22137883/)]
7. Quinn T, Johnsen S, Gale CP, Snooks H, McLean S, Woollard M, Myocardial Ischaemia National Audit Project (MINAP) Steering Group. Effects of prehospital 12-lead ECG on processes of care and mortality in acute coronary syndrome: a linked cohort study from the Myocardial Ischaemia National Audit Project. *Heart* 2014 Jun;100(12):944-950 [FREE Full text] [doi: [10.1136/heartjnl-2013-304599](https://doi.org/10.1136/heartjnl-2013-304599)] [Medline: [24732676](https://pubmed.ncbi.nlm.nih.gov/24732676/)]
8. Zègre-Hemsey JK, Asafu-Adjei J, Fernandez A, Brice J. Characteristics of prehospital electrocardiogram use in North Carolina using a novel linkage of emergency medical services and emergency department data. *Prehosp Emerg Care* 2019;23(6):772-779 [FREE Full text] [doi: [10.1080/10903127.2019.1597230](https://doi.org/10.1080/10903127.2019.1597230)] [Medline: [30885071](https://pubmed.ncbi.nlm.nih.gov/30885071/)]
9. Bosson N, Kaji AH, Niemann JT, Squire B, Eckstein M, French WJ, et al. The utility of prehospital ECG transmission in a large EMS system. *Prehosp Emerg Care* 2015;19(4):496-503. [doi: [10.3109/10903127.2015.1005260](https://doi.org/10.3109/10903127.2015.1005260)] [Medline: [25901583](https://pubmed.ncbi.nlm.nih.gov/25901583/)]

10. Sowizdraniuk J, Smereka J, Ladny J, Kaserer A, Palimonka K, Ruetzler K, et al. ECG pre-hospital teletransmission by emergency teams staffed with an emergency physician and paramedics and its impact on transportation and hospital admission. *Medicine (Baltimore)* 2019 Aug;98(34):e16636. [doi: [10.1097/MD.00000000000016636](https://doi.org/10.1097/MD.00000000000016636)] [Medline: [31441838](https://pubmed.ncbi.nlm.nih.gov/31441838/)]
11. Larochelle N, O'Keefe M, Wolfson D, Freeman K. Cellular technology improves transmission success of pre-hospital electrocardiograms. *Am J Emerg Med* 2013 Nov;31(11):1564-1570 [FREE Full text] [doi: [10.1016/j.ajem.2013.07.032](https://doi.org/10.1016/j.ajem.2013.07.032)] [Medline: [24075803](https://pubmed.ncbi.nlm.nih.gov/24075803/)]
12. Gonzalez M, Satler L, Rodrigo M, Gaglia MJ, Ben-Dor I, Maluenda G, et al. Cellular video-phone assisted transmission and interpretation of prehospital 12-lead electrocardiogram in acute st-segment elevation myocardial infarction. *J Interv Cardiol* 2011 Apr;24(2):112-118. [doi: [10.1111/j.1540-8183.2010.00609.x](https://doi.org/10.1111/j.1540-8183.2010.00609.x)] [Medline: [21457325](https://pubmed.ncbi.nlm.nih.gov/21457325/)]
13. Bansal A, Joshi R. Portable out-of-hospital electrocardiography: A review of current technologies. *J Arrhythm* 2018 Apr;34(2):129-138. [doi: [10.1002/joa3.12035](https://doi.org/10.1002/joa3.12035)] [Medline: [29657588](https://pubmed.ncbi.nlm.nih.gov/29657588/)]
14. Steijlen AS, Jansen KM, Albayrak A, Verschure DO, Van Wijk DF. A novel 12-lead electrocardiographic system for home use: development and usability testing. *JMIR Mhealth Uhealth* 2018 Jul 30;6(7):e10126 [FREE Full text] [doi: [10.2196/10126](https://doi.org/10.2196/10126)] [Medline: [30061094](https://pubmed.ncbi.nlm.nih.gov/30061094/)]
15. Alis C, del Rosario C, Buenaobra B, Mar Blanca C. Lifelink: 3G-based mobile telemedicine system. *Telemed J E Health* 2009 Apr;15(3):241-247. [doi: [10.1089/tmj.2008.0098](https://doi.org/10.1089/tmj.2008.0098)] [Medline: [19382861](https://pubmed.ncbi.nlm.nih.gov/19382861/)]
16. Cha YD, Yoon G. Ubiquitous health monitoring system for multiple users using a ZigBee and WLAN dual-network. *Telemed J E Health* 2009 Nov;15(9):891-897. [doi: [10.1089/tmj.2009.0049](https://doi.org/10.1089/tmj.2009.0049)] [Medline: [19919196](https://pubmed.ncbi.nlm.nih.gov/19919196/)]
17. Orlov OI, Drozdov DV, Doarn CR, Merrell RC. Wireless ECG monitoring by telephone. *Telemed J E Health* 2001;7(1):33-38. [doi: [10.1089/153056201300093877](https://doi.org/10.1089/153056201300093877)] [Medline: [11321707](https://pubmed.ncbi.nlm.nih.gov/11321707/)]
18. Woodward B, Istepanian R, Richards C. Design of a telemedicine system using a mobile telephone. *IEEE Trans Inf Technol Biomed* 2001 Mar;5(1):13-15. [doi: [10.1109/4233.908361](https://doi.org/10.1109/4233.908361)] [Medline: [11300210](https://pubmed.ncbi.nlm.nih.gov/11300210/)]
19. Zhao X, Fei D, Doarn CR, Harnett B, Merrell R. A telemedicine system for wireless home healthcare based on Bluetooth and the Internet. *Telemed J E Health* 2004;10(Suppl 2):S-110-6. [Medline: [23570221](https://pubmed.ncbi.nlm.nih.gov/23570221/)]
20. G\*Power: Statistical Power Analyses for Windows and Mac. Heinrich Heine Universität Düsseldorf. URL: <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html> [accessed 2020-12-22]
21. Brooke J. SUS-a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation in Industry*. London: Taylor & Francis; Jun 11, 1996:189-194.
22. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 2009 May;4(3):114-123 [FREE Full text]
23. Barquera S, Pedroza-Tobías A, Medina C, Hernández-Barrera L, Bibbins-Domingo K, Lozano R, et al. Global overview of the epidemiology of atherosclerotic cardiovascular disease. *Arch Med Res* 2015 Jul;46(5):328-338. [doi: [10.1016/j.arcmed.2015.06.006](https://doi.org/10.1016/j.arcmed.2015.06.006)] [Medline: [26135634](https://pubmed.ncbi.nlm.nih.gov/26135634/)]
24. AlGhatrif M, Lindsay J. A brief review: history to understand fundamentals of electrocardiography. *J Community Hosp Intern Med Perspect* 2012;2(1):14383 [FREE Full text] [doi: [10.3402/jchimp.v2i1.14383](https://doi.org/10.3402/jchimp.v2i1.14383)] [Medline: [23882360](https://pubmed.ncbi.nlm.nih.gov/23882360/)]
25. Grim P, Feldman T, Martin M, Donovan R, Nevins V, Childers RW. Cellular telephone transmission of 12-lead electrocardiograms from ambulance to hospital. *Am J Cardiol* 1987 Sep 15;60(8):715-720. [doi: [10.1016/0002-9149\(87\)90388-2](https://doi.org/10.1016/0002-9149(87)90388-2)] [Medline: [3661440](https://pubmed.ncbi.nlm.nih.gov/3661440/)]
26. Aufderheide TP, Hendley GE, Thakur R, Mateer JR, Stueven HA, Olson DW, et al. The diagnostic impact of prehospital 12-lead electrocardiography. *Ann Emerg Med* 1990 Nov;19(11):1280-1287. [doi: [10.1016/s0196-0644\(05\)82288-7](https://doi.org/10.1016/s0196-0644(05)82288-7)] [Medline: [2240725](https://pubmed.ncbi.nlm.nih.gov/2240725/)]
27. Hsieh J, Lin B, Wu F, Chang P, Tsuei Y, Yang C. Ambulance 12-lead electrocardiography transmission via cell phone technology to cardiologists. *Telemed J E Health* 2010 Oct;16(8):910-915. [doi: [10.1089/tmj.2010.0034](https://doi.org/10.1089/tmj.2010.0034)] [Medline: [20925562](https://pubmed.ncbi.nlm.nih.gov/20925562/)]
28. Werman HA, Newland R, Cotton B. Transmission of 12-lead electrocardiographic tracings by emergency medical technician-basics and emergency medical technician-intermediates: a feasibility study. *Am J Emerg Med* 2011 May;29(4):437-440. [doi: [10.1016/j.ajem.2010.01.015](https://doi.org/10.1016/j.ajem.2010.01.015)] [Medline: [20825850](https://pubmed.ncbi.nlm.nih.gov/20825850/)]
29. Garvey JL, MacLeod BA, Sopko G, Hand MM, National Heart Attack Alert Program (NHAAP) Coordinating Committee, National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health. Pre-hospital 12-lead electrocardiography programs: a call for implementation by emergency medical services systems providing advanced life support--National Heart Attack Alert Program (NHAAP) Coordinating Committee; National Heart, Lung, and Blood Institute (NHLBI); National Institutes of Health. *J Am Coll Cardiol* 2006 Feb 07;47(3):485-491 [FREE Full text] [doi: [10.1016/j.jacc.2005.08.072](https://doi.org/10.1016/j.jacc.2005.08.072)] [Medline: [16458125](https://pubmed.ncbi.nlm.nih.gov/16458125/)]
30. Ouellette L, VanDePol E, Chassee T, Tavares E, Flannigan M, Jones J. Emergency department electrocardiogram images sent through the mobile phone: feasibility and accuracy. *Am J Emerg Med* 2018 Apr;36(4):731-732. [doi: [10.1016/j.ajem.2017.09.010](https://doi.org/10.1016/j.ajem.2017.09.010)] [Medline: [28917438](https://pubmed.ncbi.nlm.nih.gov/28917438/)]
31. Faxon DP. Early reperfusion strategies after acute ST-segment elevation myocardial infarction: the importance of timing. *Nat Clin Pract Cardiovasc Med* 2005 Jan;2(1):22-28. [doi: [10.1038/ncpcardio0065](https://doi.org/10.1038/ncpcardio0065)] [Medline: [16265339](https://pubmed.ncbi.nlm.nih.gov/16265339/)]
32. Nissen SE. Saving lives in acute ST-elevation myocardial infarction: the door-to-balloon initiative. *Curr Cardiol Rep* 2007 Apr;9(2):79-81. [doi: [10.1007/BF02938331](https://doi.org/10.1007/BF02938331)] [Medline: [17430673](https://pubmed.ncbi.nlm.nih.gov/17430673/)]



33. Bassand J, Danchin N, Filippatos G, Gitt A, Hamm C, Silber S, et al. Implementation of reperfusion therapy in acute myocardial infarction. A policy statement from the European Society of Cardiology. *Eur Heart J* 2005 Dec;26(24):2733-2741. [doi: [10.1093/eurheartj/ehi673](https://doi.org/10.1093/eurheartj/ehi673)] [Medline: [16311237](https://pubmed.ncbi.nlm.nih.gov/16311237/)]

## Abbreviations

**ACS:** acute coronary syndrome  
**C-ECG:** conventional electrocardiogram  
**CVD:** cardiovascular disease  
**ECG:** electrocardiogram  
**EMS:** emergency medical service  
**EMT:** emergency medical technician  
**LTE:** long-term evolution  
**P12ECG:** prehospital 12-lead electrocardiogram  
**P-ECG:** patchy-type wireless electrocardiogram  
**STEMI:** ST-segment elevation myocardial infarction  
**SUS:** System Usability Scale

*Edited by L Buis; submitted 05.09.20; peer-reviewed by KC Wong; comments to author 02.10.20; revised version received 22.12.20; accepted 11.03.21; published 01.04.21.*

*Please cite as:*

*Yoon S, Kim T, Roh T, Chang H, Hwang SY, Yoon H, Shin TG, Sim MS, Jo JJ, Cha WC*

*Twelve-Lead Electrocardiogram Acquisition With a Patchy-Type Wireless Device in Ambulance Transport: Simulation-Based Randomized Controlled Trial*

*JMIR Mhealth Uhealth* 2021;9(4):e24142

URL: <https://mhealth.jmir.org/2021/4/e24142>

doi: [10.2196/24142](https://doi.org/10.2196/24142)

PMID: [33792550](https://pubmed.ncbi.nlm.nih.gov/33792550/)

©Sunyoung Yoon, Taerim Kim, Taehwan Roh, Hansol Chang, Sung Yeon Hwang, Hee Yoon, Tae Gun Shin, Min Seob Sim, Ik Joon Jo, Won Chul Cha. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 01.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Force-Sensitive Mat for Vertical Jump Measurement to Assess Lower Limb Strength: Validity and Reliability Study

Erik Vanegas<sup>1\*</sup>, MSc; Yolocuauhtli Salazar<sup>2\*</sup>, PhD; Raúl Igual<sup>1\*</sup>, PhD; Inmaculada Plaza<sup>1\*</sup>, PhD

<sup>1</sup>Electrical/Electronics Engineering and Communications Department, EUP Teruel, Universidad de Zaragoza, Teruel, Spain

<sup>2</sup>Tecnológico Nacional de México, IT Durango, Durango, Mexico

\* all authors contributed equally

**Corresponding Author:**

Erik Vanegas, MSc

Electrical/Electronics Engineering and Communications Department, EUP Teruel

Universidad de Zaragoza

Atarazana 2

Teruel, 44003

Spain

Phone: 34 978618102

Email: [erikvanegas599@gmail.com](mailto:erikvanegas599@gmail.com)

## Abstract

**Background:** Vertical jump height is widely used in health care and sports fields to assess muscle strength and power from lower limb muscle groups. Different approaches have been proposed for vertical jump height measurement. Some commonly used approaches need no sensor at all; however, these methods tend to overestimate the height reached by the subjects. There are also novel systems using different kind of sensors like force-sensitive resistors, capacitive sensors, and inertial measurement units, among others, to achieve more accurate measurements.

**Objective:** The objective of this study is twofold. The first objective is to validate the functioning of a developed low-cost system able to measure vertical jump height. The second objective is to assess the effects on obtained measurements when the sampling frequency of the system is modified.

**Methods:** The system developed in this study consists of a matrix of force-sensitive resistor sensors embedded in a mat with electronics that allow a full scan of the mat. This mat detects pressure exerted on it. The system calculates the jump height by using the flight-time formula, and the result is sent through Bluetooth to any mobile device or PC. Two different experiments were performed. In the first experiment, a total of 38 volunteers participated with the objective of validating the performance of the system against a high-speed camera used as reference (120 fps). In the second experiment, a total of 15 volunteers participated. Raw data were obtained in order to assess the effects of different sampling frequencies on the performance of the system with the same reference device. Different sampling frequencies were obtained by performing offline downsampling of the raw data. In both experiments, countermovement jump and countermovement jump with arm swing techniques were performed.

**Results:** In the first experiment an overall mean relative error (MRE) of 1.98% and a mean absolute error of 0.38 cm were obtained. Bland-Altman and correlation analyses were performed, obtaining a coefficient of determination equal to  $R^2=0.996$ . In the second experiment, sampling frequencies of 200 Hz, 100 Hz, and 66.6 Hz show similar performance with MRE below 3%. Slower sampling frequencies show an exponential increase in MRE. On both experiments, when dividing jump trials in different heights reached, a decrease in MRE with higher height trials suggests that the precision of the proposed system increases as height reached increases.

**Conclusions:** In the first experiment, we concluded that results between the proposed system and the reference are systematically the same. In the second experiment, the relevance of a sufficiently high sampling frequency is emphasized, especially for jump trials whose height is below 10 cm. For trials with heights above 30 cm, MRE decreases in general for all sampling frequencies, suggesting that at higher heights reached, the impact of high sampling frequencies is lesser.

(*JMIR Mhealth Uhealth* 2021;9(4):e27336) doi:[10.2196/27336](https://doi.org/10.2196/27336)

**KEYWORDS**

vertical jump; mHealth; mobile health; force-sensitive resistor; lower limb strength; leg strength

## Introduction

Vertical jump height is one of the physical skills commonly used to assess overall performance in human beings, and more specifically, it is used to assess performance and muscle power of the quadriceps, hamstrings, and gastrocnemius muscle groups in the lower limbs [1,2]. Measurement of the performance of this skill is commonly performed on athletes in sports like basketball [3,4], football [5], netball [6], swimming [7], and others. This skill performance can also provide important data from people with no relevant sports past.

In the literature, there are many protocols to prove or validate the proposed systems. Among the different kind of jumps performed in those protocols, there are jumps with and without countermovement [1,4,5,8-15], jumps with and without arm swing [12,16], drop jumps [1,8,17], single and double leg jump [6], continuous jumps [4,17], squat jumps [1,2,4,12], and loaded squat jumps [7]. With any of these types of jumps, height reached by the user can be analyzed, but the jumps most commonly used in all related work are the countermovement and squat jumps.

Kibele [15] and Moir [18] reported that irregularities detected in measurements of any kind of jump execution may be linked to changes in the posture of a subject during flight due to change in the center of mass of the subject during the jump. Bui et al [13] found that some common errors obtained during measurement were caused by body movements like knee, hip, and ankle bending during flight time and landing. Also, Aragon-Vargas [19] states that ascending and descending phases of flight time must be of the same length of time, but in his work descending time was significantly longer, suggesting that participants descended with their bodies partially crouched.

There are several techniques to measure jump height, each of which uses a different kind of sensor or no sensors at all. Methods like the Sargent jump [13] and Vertec device [6,9,10] require no sensors and often are used as reference measurements. However, these methods often show overestimation on jump height, and this could be due to arm stretching performed unconsciously by the user. Among systems developed in the literature, different kind of sensors are used like force-sensitive resistors (FSRs) [3,16], capacitive sensors [5], inertial measurement units [2,4,8,10,17], electromyography sensors [1,6], kinematic sensors [6], ultrasonic sensors [20], microswitches [9], video cameras, [11] and optical sensors [4,12].

The studies of Drazan et al [3] and Boukhenous and Attari [16] are most closely related to our work, as both of their systems also use resistive sensors. However, only one sensor is used for the whole sensing area. Drazan et al [3] proposed a system based on a single FSR sensor whose total sensing area is around 3 cm<sup>2</sup>, with an Arduino board as microcontroller. This system calculated jump height through the flight-time formula, by measuring the time the FSR sensor is not detecting any pressure. In the work presented by Boukhenous and Attari [16], two metallic strain gauges were placed in the center of a rigid platform to measure the force applied by the ground. In this case, vertical ground reaction forces were used to calculate jump

height. Rico et al [2] used pressure sensors located at the forefoot of the user to calculate flight time during vertical jump and compare it with data obtained from an inertial measurement unit system. However, few or no specific number of subjects are used in these studies, and no specific protocols or jump trials are performed. Also, no reference systems are used to compare the performance of the developed systems.

Camera-based systems are used as reference systems or as proposed methods to validate. In some studies, the famous motion capture system commonly used in videogames is used as a reference device [5,17,21]. Other studies use a similar method by tracking body markers placed strategically on the body [4,14,21]. Balsalobre-Fernandez et al [11,22] have analyzed the effectiveness and reliability of high-speed cameras as methods to estimate vertical jump height. In these studies, flight time of the subject is calculated by selecting the takeoff and landing frames of the recorded videos of jump trials, and, by applying the flight-time formula, jump height is obtained [3,8,9,11,14,16,22].

Only a few studies perform a validation with a relatively high number of subjects. Some studies that fulfill this criterion are the ones presented by Nuzzo et al [10], Casartelli et al [4], Glatthorn et al [12], Moir [18], and Aragon-Vargas [19]. However, these studies compare different commercially available devices (contact mat, force plate, cameras, etc), and no novel system is developed by the researchers. Bui et al [13] fulfills the criterion and proposes a novel optical system whose performance is compared against commercial devices. This system calculates jump height through the flight-time formula.

This study presents a newly developed low-cost system for measuring height reached by users during vertical jump comprising a matrix of FSR sensors embedded on a mat. The height of the vertical jump is calculated through the flight-time formula [3,8,9,14,16]. One advantage this system offers against other pressure-sensitive systems [2,3,16] is a higher real sensing area, higher resolution, and higher precision, as this system works with 256 FSR sensors distributed around the mat in 16 rows and 16 columns, in comparison with the other systems that use a single sensor. The total sensing area, dimensions of FSR sensors matrix, and each individual FSR sensor area can be modified on different versions of the proposed mat. Also, as this system is environmental, it needs no adjustment regardless of the physical characteristics of the user like body type, weight, height, or foot size [2]. Another advantage of this system is that the calculated vertical jump height is directly sent to a PC or mobile device of the health care professional's choice, unlike other methods that require postprocessing analysis, as in the high-speed camera method. The main objective of this study is to validate the reliability of the proposed system for future clinical studies.

## Methods

### System Construction

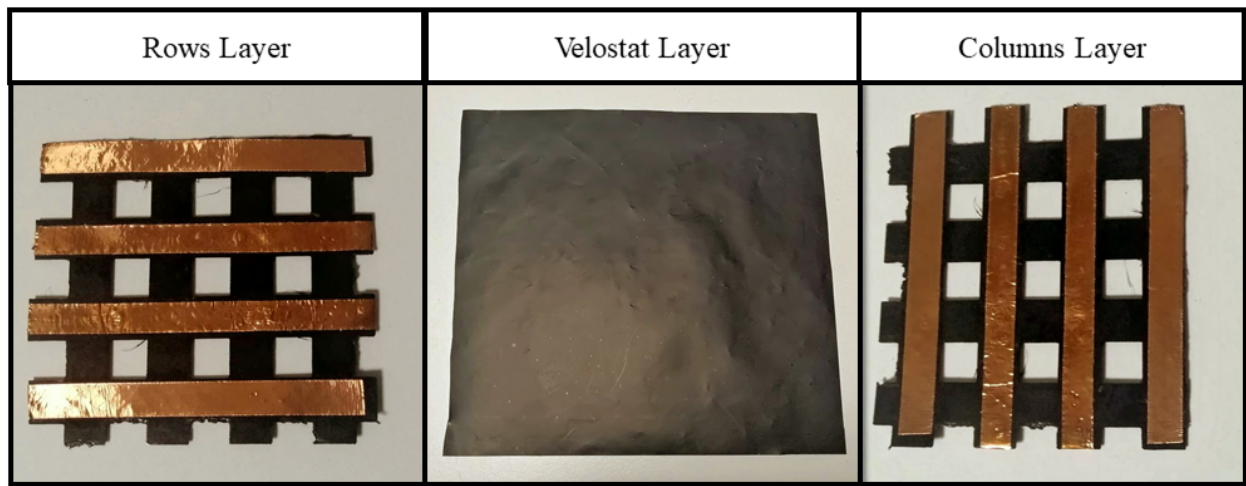
The proposed system consists of 2 parts: a resistive pressure-sensitive mat constructed with an FSR sensor array and the electronic system. The mat is composed of 3 layers.

One layer contains thin and flat copper wires distributed in a column arrangement along a flexible 3D printed grid; another layer is composed of the same copper wires but distributed in a row arrangement on another flexible 3D printed grid. A third layer is placed between the layers consisting of Velostat material, a pressure-sensitive material that behaves as a resistor whose value drops whenever pressure is exerted upon it. In this way, variations on the resistance values on every intersection of rows and columns when pressure is exerted over the mat can be measured. Some typical characteristic problems with Velostat material are repeatability, nonlinearity, and hysteresis [23,24]. For this application, however, these features are not relevant

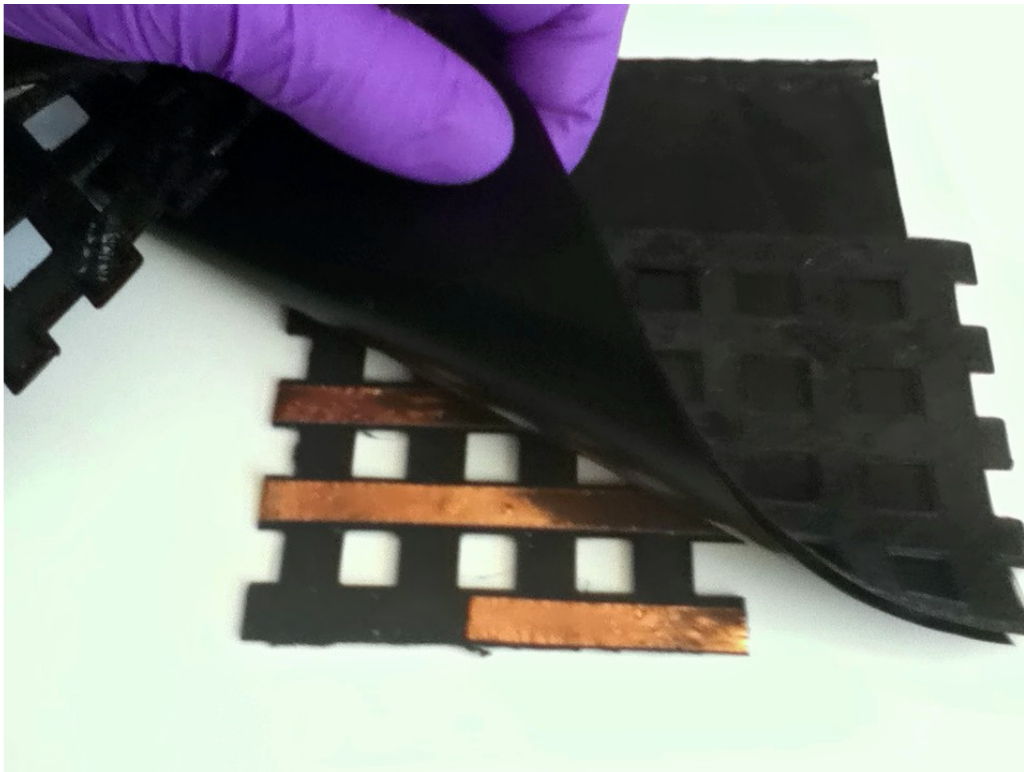
because high precision is not needed; the mat only needs to be able to detect a heavy body placed on it (average human body weight). More information about the development of this mat is documented on the work of Medrano et al [25].

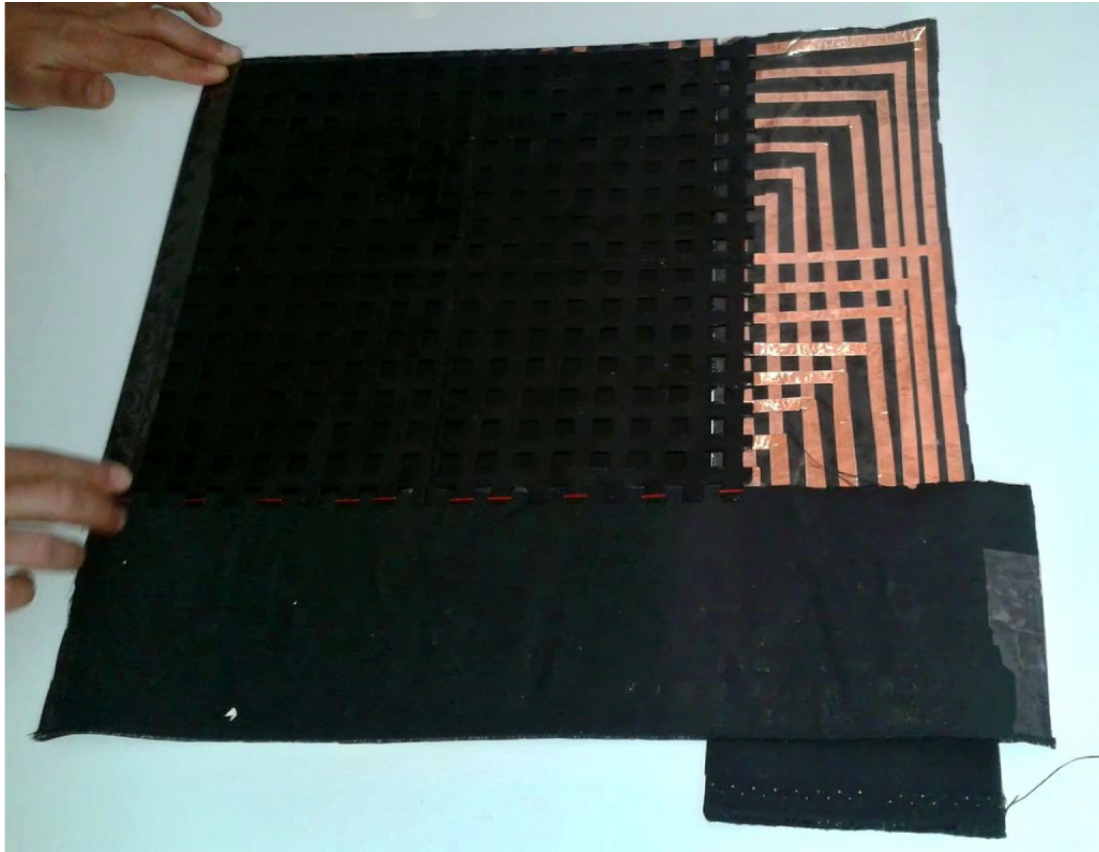
Figure 1 shows the different layers of an FSR matrix with smaller dimensions (4 rows and 4 columns). Figure 2 shows an example of the placement of the overlapped layers. The total sensing area of the mat used for this study is 30×30 cm, with 16 rows and 16 columns, 1 cm width each. This way, the area of each of the individual FSR sensors is equal to 1 cm<sup>2</sup>. In Figure 3, a developed mat is shown.

**Figure 1.** Different layers comprising a force-sensitive resistor matrix with smaller dimensions.



**Figure 2.** Example of a smaller size matrix and how layers are placed.



**Figure 3.** Developed force-sensitive resistor sensor mat.

Due to the number of operations needed for a full scan of the mat, a high frequency microprocessor must be used for data processing, as the time complexity of these operations grows in exponential order. The STM32F103C8T6 microprocessor (STMicroelectronics) was selected due to its 72 MHz CPU frequency, with which a sampling frequency of 200 Hz is achieved. Other microprocessors with lower CPU frequencies (like the ATmega328P, Microchip Technology Inc) would not achieve the desired sampling frequency. Also two 16-1 multiplexers 74HC4067 are needed for an efficient scan process of the whole mat. For data transmission, a Bluetooth HC-05 module is used. Bluetooth technology was selected due to its ease of connection with different devices, especially with smartphones and tablets, which offers health care professionals the choice of an easy-to-transport monitoring device. Other electronic elements included in the system are a TP4056 battery-charging module and a Lipo battery of 3.7 V and 150 mAh capacity, allowing continuous functioning of the system for up to 2 hrs. A block diagram of the proposed system is shown in [Figure 4](#).

The algorithm used for this system consists of calculating the summation of every FSR sensor of the mat. For each FSR sensor, the voltage value obtained by the analog-to-digital converter of the microcontroller is given in bits (from 0 to 4095), and this resulting value is used for the calculations. A threshold is used for the system to decide whether a person is standing on the mat or not. To calculate an appropriate value for this threshold, data were collected from 16 volunteers (5 female and

11 male), with an average weight of 74.81 (SD 15.25) kg and foot size of 26.93 (SD 1.94) cm. The volunteers were asked to stand on the mat barefoot in 4 different positions: with both feet standing still and on their forefoot and with one foot standing still and on their forefoot. Maximum values of pressure exerted on FSR sensors were used as reference for normalization, and the minimum value for activation of FSR sensors was considered as no volunteer standing on the mat.

Using such criteria, on average when standing still over the mat with both feet subjects activated 71.66% of the FSR sensors, and when standing on their forefoot with both feet, 28.9% of the FSR sensors were activated. When standing still and on their forefoot with only one foot an activation of 40.37% and 18.40% of the FSR sensors was registered, respectively. The minimum value of FSR sensors activation is registered when standing on one foot on their forefoot, with a value of 12.09%. All results are summarized in [Table 1](#). By taking these results into account, and if it is assumed that volunteers may land first with one foot on their forefoot after a jump, a proper threshold should be proposed below the minimum value of FSR sensors activation. For this study, a threshold of 9% of FSR sensors activation is used. This threshold was chosen to be at three-quarters between the zero FSR sensor activation and the minimum FSR sensor activation recorded, to avoid any misreading from mechanical oscillation. It is worth noting that the minimum recorded value from sensor activation is an outlier. In future studies, the possibility of adding a personalized threshold for every subject could be assessed.

Figure 4. Block diagram of the proposed system.

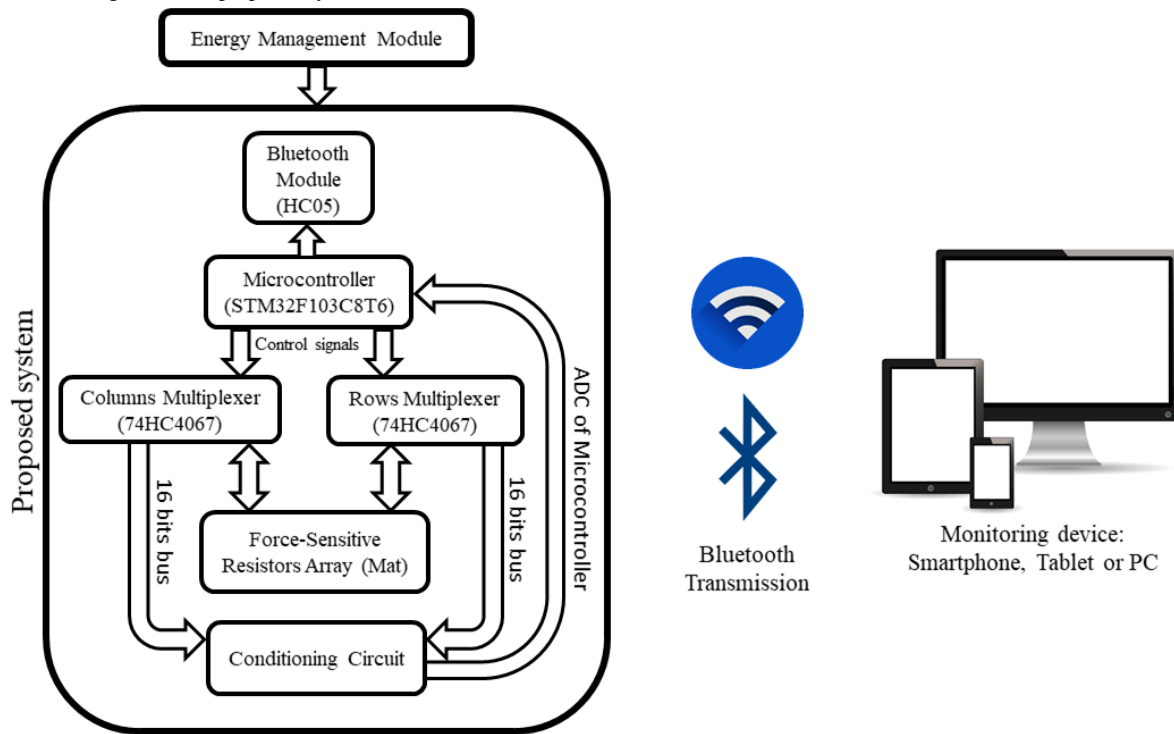


Table 1. Normalized force-sensitive resistor sensors activation registered from volunteers standing at different positions on the mat; standing still and on their toes with both feet and standing still and on their toes with one foot.

Sensor activation	Both feet %		One foot %	
	Standing	Toes	Standing	Toes
Average	71.66	28.90	40.37	18.40
Maximum	100.00	43.31	61.45	27.00
Minimum	44.00	23.70	22.64	12.09

For every position sensor activation percentage, a correlation analysis with the weight and foot size of the volunteers was performed. Analysis suggests that sensor activation is only moderately impacted by the weight of volunteers, and foot size of volunteers has a low impact on sensor activation. In Table 2, correlation values for different positions analyzed on the mat are shown.

To calculate the height reached by the user during the vertical jump, when the subject jumps and there is no contact on the

mat, the system counts the elapsed time until the subject lands on the mat again (flight time), and with the calculated time the flight-time formula is used [3,8,9,14,16] to predict the height reached. This formula is defined as:  $Height = g\Delta t^2/8$ , where  $g$  is the constant value of gravity force  $g=9.81 \text{ m/s}^2$  and  $\Delta t$  is the flight time obtained by the system. Once the height of the vertical jump is obtained, this value is wirelessly sent via Bluetooth to the monitoring device selected by the health care professional.

Table 2. Pearson correlation coefficient values (R values) for different positions analyzed, calculated for weight and foot size of volunteers.

Position	Weight	Foot size
<b>Both feet</b>		
Standing	-.684	-.522
Forefoot	-.411	-.241
<b>One foot</b>		
Standing	-.447	-.397
Forefoot	-.394	-.463

## Experimental Setup

Two different experiments were performed. The purpose of the first experiment, in which 38 volunteers participated, was to validate the reliability of the proposed system. For the second experiment, the objective was to compare the effects of different sampling frequencies when calculating the height reached on the jumps, and 15 volunteers participated. The protocol used for each experiment is the same. For the first experiment, data are directly processed by the microcontroller and the predicted value is sent to the selected monitoring station. In the second experiment, raw data are obtained to perform an offline downsampling to analyze the effects of different sampling frequencies on the predicted result.

Researchers asked for assistance from a sports and fitness center to recruit volunteers who attended the center regularly for physical training. Researchers visited this center with all the necessary equipment for the implementation of the protocols, and installed it in an area specified by the managers of the center. No specific physical attributes were required from the volunteers, as these characteristics should not affect the performance of the system. Researchers approached people at the center, explained the purpose of the study, and politely asked for their collaboration on the protocols if they were available at any given moment. Before starting any trial, volunteers were asked if they had any kind of injury that could affect their physical integrity when performing the protocol, and if so, the trials would not proceed. Every volunteer gave their written consent for the performance of the proposed protocol. The

countermovement jump (CMJ) and countermovement jump with arm swing (CMJAS) techniques were selected for this protocol. These jumping techniques are commonly used as a measure to assess the overall force and explosive power of the lower body muscles on a person [26], and it is considered as the most reliable jumping test for this purpose [27]. By adding an arm swing to the CMJ, with the proper technique, the height reached by the person is increased around one-third and up to two-thirds [28-30], which increases the dynamic range of data obtained.

In the proposed protocol, volunteers were asked to stand on a marker placed on the center of the mat and perform 3 medium-to-maximal effort CMJs, with their hands fixed at the waist, with 5 to 10 seconds rest between trials. This technique is depicted in Figure 5. After these jumps, the volunteers were asked to perform CMJAS this time, and following the same scheme. This technique is depicted in Figure 6. As a reference system, all trials were recorded on video with a high-speed camera (120 fps). The camera was placed 1.3 m away from the mat, perpendicular to the sagittal plane of the volunteer and 20 cm above the ground, held by a tripod. The setup for this protocol is depicted in Figure 7.

To measure height reached by the subject with the video reference, the takeoff and landing frames were selected manually like in the studies of Balsalobre-Fernandez et al [11,22], and height was calculated by using the elapsed time between the frames using the flight-time formula.

**Figure 5.** Countermovement jump technique, step by step.

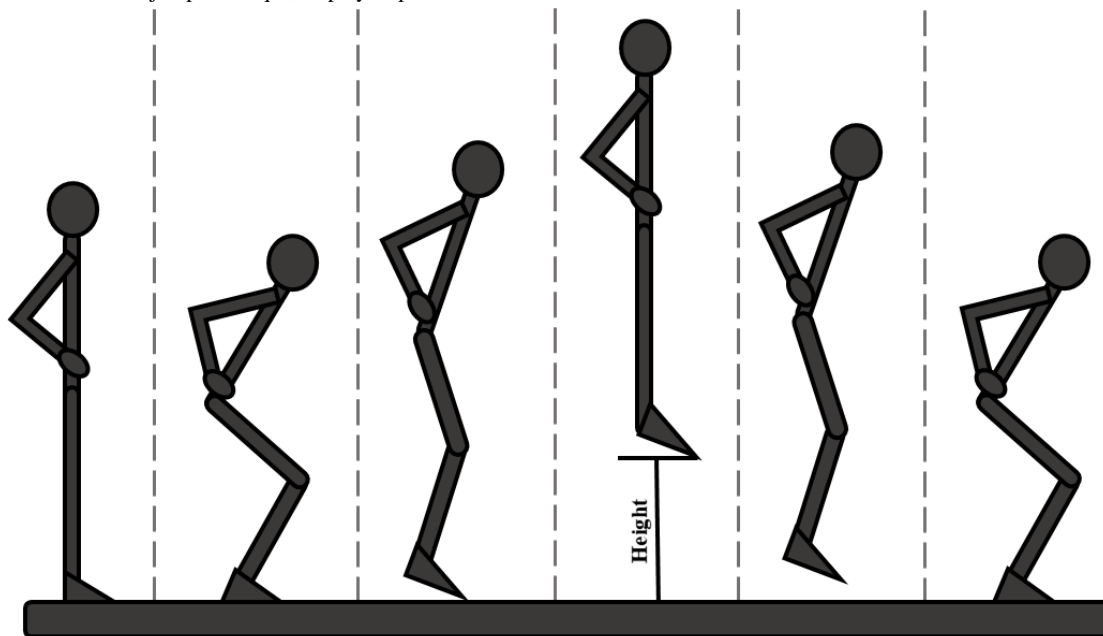


Figure 6. Countermovement jump with arm swing technique, step by step.

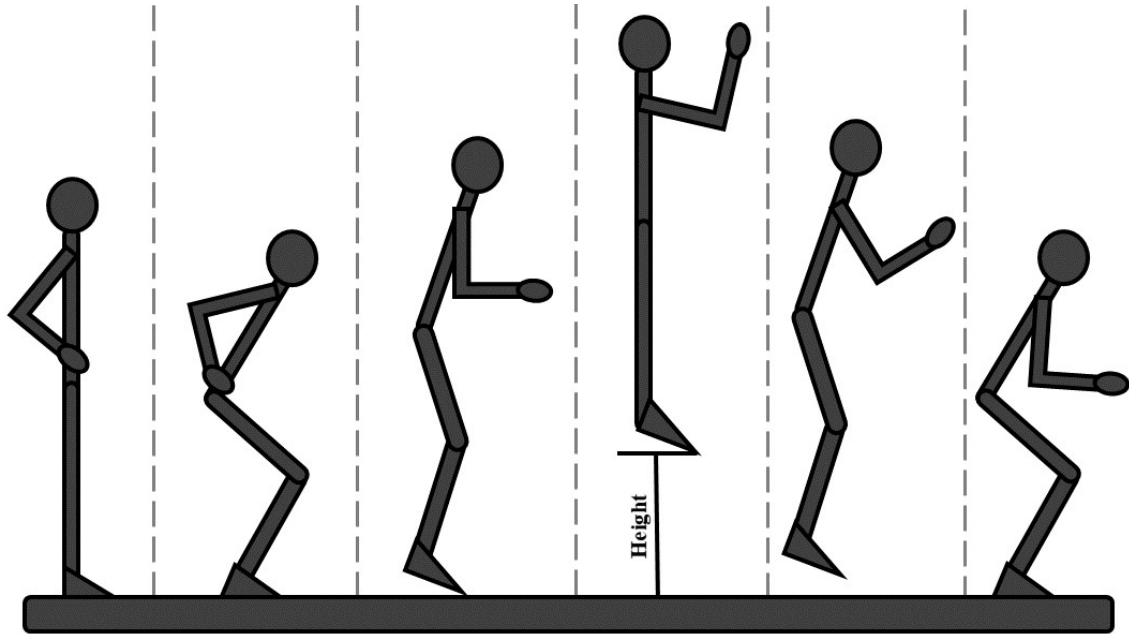
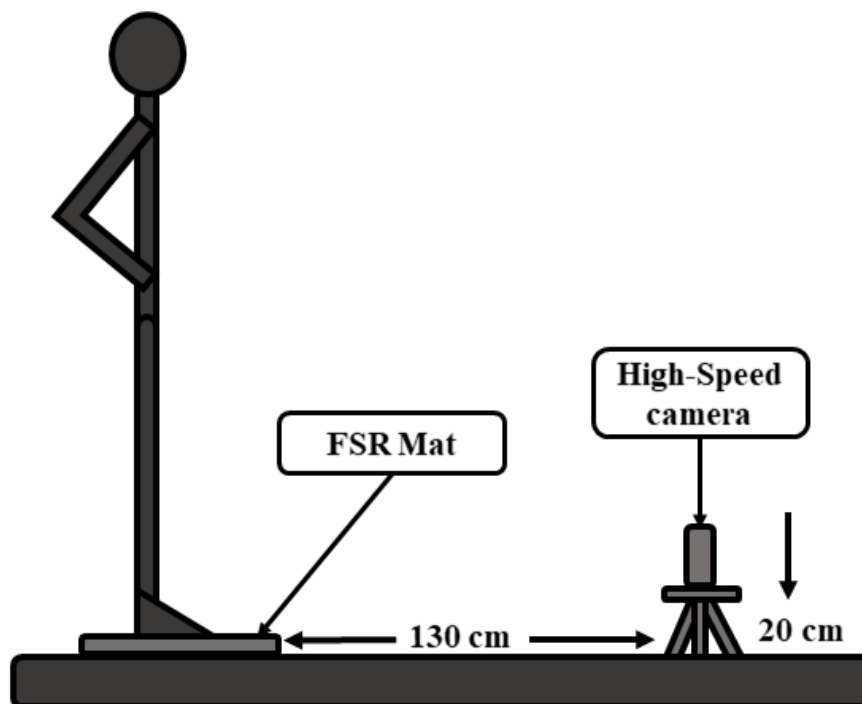


Figure 7. Experimental setup used for proposed experiments.



## Results

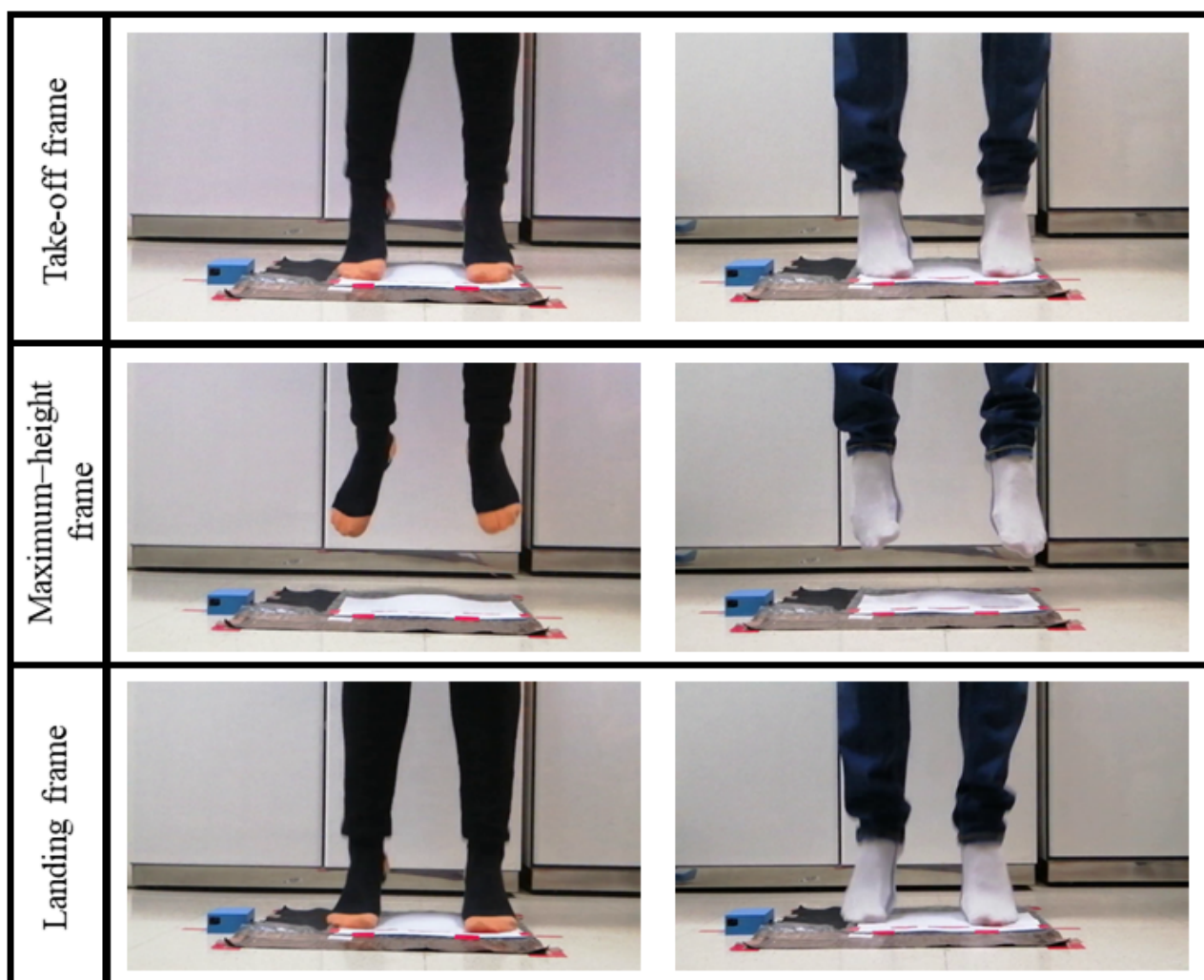
### First Experiment: System Validation

For the first experiment, a total of 228 jumps (114 CMJs and 114 CMJASs) were performed for each, the proposed system and the video reference. An example of the recorded jumps is shown in [Figure 8](#).

To analyze the proposed system performance, mean relative error (MRE) and mean absolute error (MAE) were calculated for the overall jump trials and for each technique, CMJ and CMJAS. The MRE obtained from all 228 trials was 1.98%. For CMJ and CMJAS, relative errors were 2.17% and 1.78%, respectively. MAE obtained from all jump trials was 0.38 cm, and for CMJ and CMJAS, the errors obtained were 0.34 cm and 0.42 cm, respectively. These results are summarized in [Table 3](#).



**Figure 8.** Two volunteers performing the proposed protocol showing the different phases of the jumps: takeoff frame, maximum-height frame, and landing frame.



**Table 3.** Mean absolute error and mean relative error values for overall jump trials, only countermovement jump, and only countermovement jump with arm swing trials.

Trials	MAE <sup>a</sup> (cm)	MRE <sup>b</sup> (%)
Overall	0.38	1.98
CMJ <sup>c</sup>	0.34	2.17
CMJAS <sup>d</sup>	0.42	1.78

<sup>a</sup>MAE: mean absolute error.

<sup>b</sup>MRE: mean relative error.

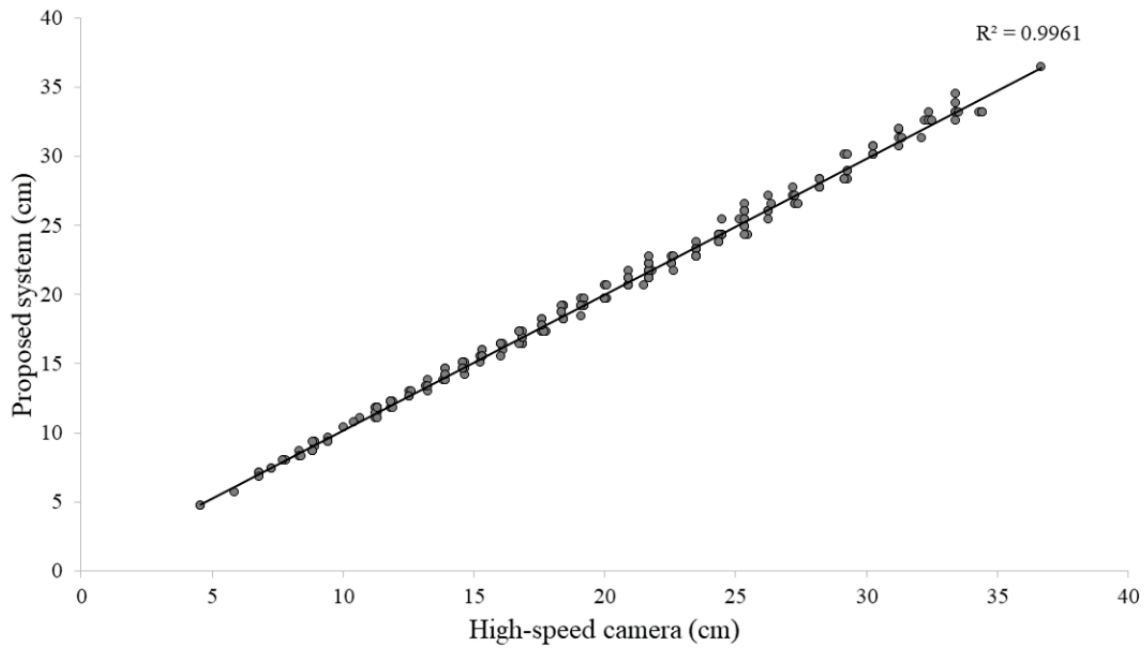
<sup>c</sup>CMJ: countermovement jump.

<sup>d</sup>CMJAS: countermovement jump with arm swing.

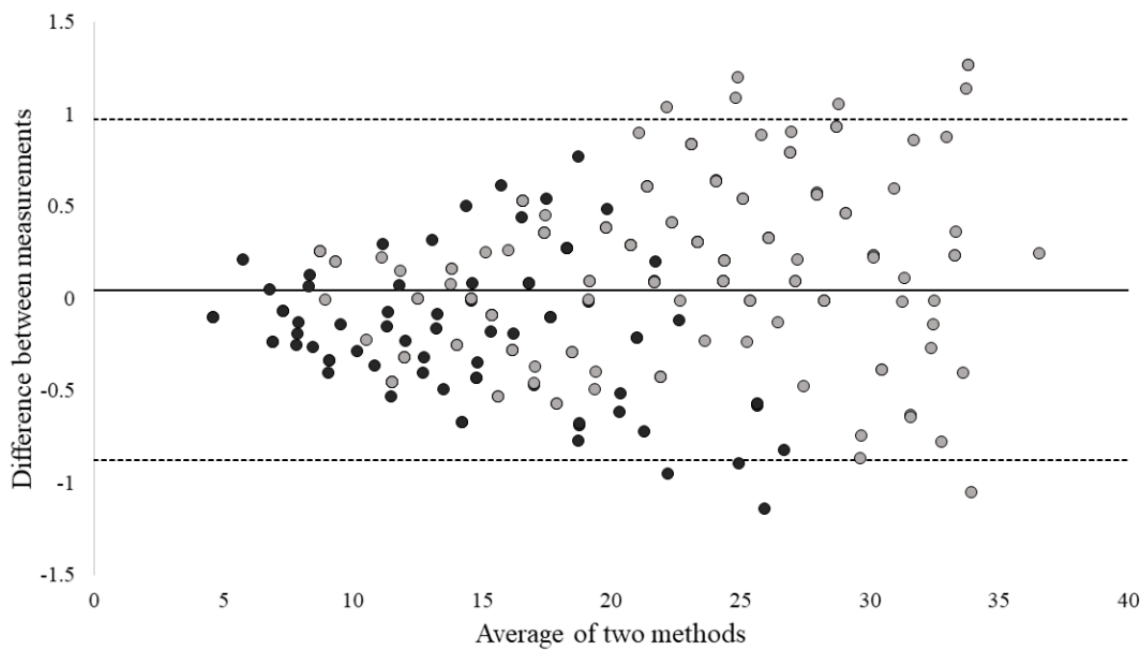
Correlation and Bland-Altman analyses were performed from data obtained and are shown in Figure 9 and Figure 10, respectively. Correlation analysis shows a coefficient of determination of  $R^2=.996$ . These analyses demonstrate that the

proposed system not only has a high correlation, but it shows that the difference of the two paired measurements is really low, which means that both methods produce systematically the same results.

**Figure 9.** Correlation graph comparing both measuring methods for the first experiment, showing a coefficient of determination of  $R^2=0.996$ .



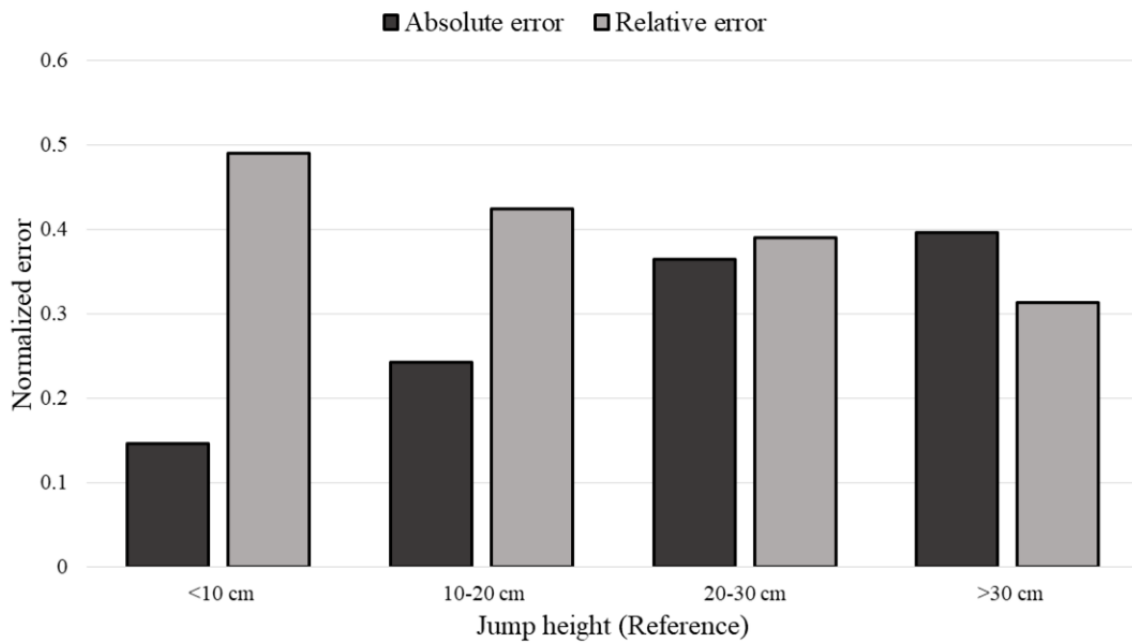
**Figure 10.** Bland-Altman plot of both measuring methods: countermovement jump depicted by dark gray points and countermovement jump with arm swing depicted by light gray points.



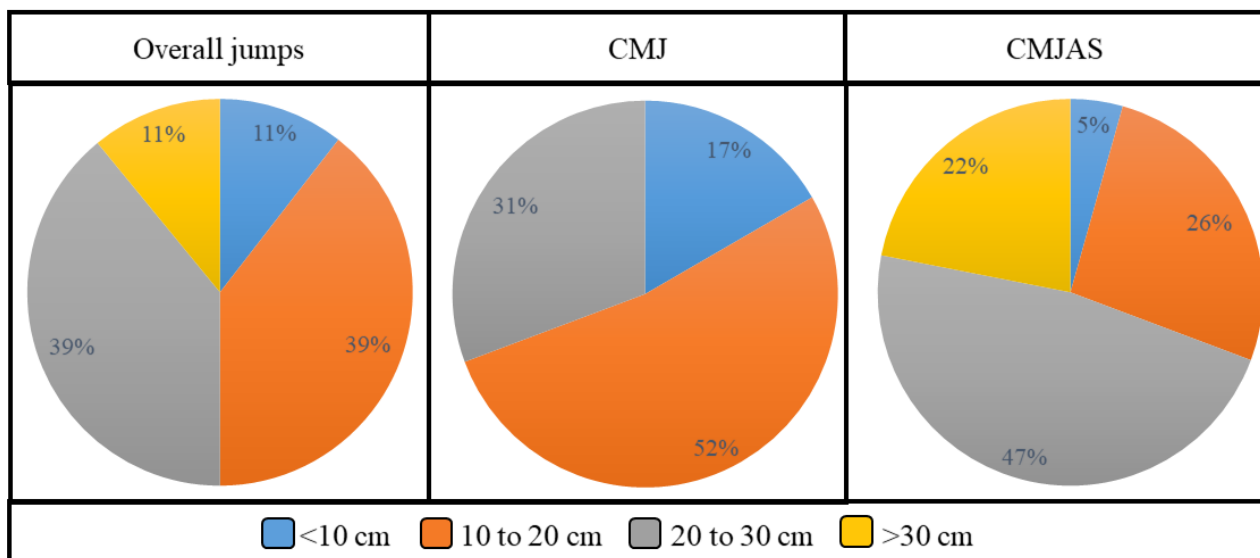
In Figure 11, the normalized MAE and MRE are shown for different ranges of jump heights reached. By analyzing the different ranges of height reached by the volunteers, which are <10 cm, 10 to 20 cm, 20 to 30 cm, and >30 cm, MREs obtained were 2.38%, 2.07%, 1.90%, and 1.54%, respectively. MAE obtained were 0.18 cm, 0.31 cm, 0.46 cm, and 0.50 cm, respectively. From this data, no significant difference can be found. However, it can be noticed that MAE increases as jump height increases, while MRE decreases.

Figure 12 shows the charts with distribution of the heights reached by the volunteers when performing the jump trials. For CMJ, no volunteer was able to surpass the 30 cm height. However, by adding the arm swing, 22% of the volunteers surpassed the 30 cm height. Also, for CMJAS trials, 47% of the subjects reached a height ranging from 20 to 30 cm, compared with CMJ, in which only 31% of the subjects reached this height.

**Figure 11.** Normalized mean absolute error and mean relative error, divided in different ranges of height reached during vertical jump.



**Figure 12.** Distribution of different height ranges reached by the users. Overall trials, only countermovement jump trials, and only countermovement jump with arm swing trials are shown.



**Second Experiment: Sampling Frequencies Comparison**

In this experiment, the effects of different sampling frequencies were analyzed. Raw data from the system was obtained for a total of 90 jumps (45 CMJs and 45 CMJASs). An offline emulation of different sampling frequencies was performed through downsampling of this raw data. This means samples are removed to emulate a slower sampling frequency. With this method, and with the base sampling period of 5 ms from the system, 200 Hz, 100 Hz, 66.6 Hz, 50 Hz, 40 Hz, 33.3 Hz, 28.5 Hz, 25 Hz, 22.2 Hz, and 20 Hz frequencies were emulated.

Similar to the first experiment, the error was calculated using the high-speed camera as reference. For this analysis, only MRE was obtained for each sampling frequency to assess which frequencies are able to maintain a relative error below 5%.

Results show that sampling frequencies of 200 Hz, 100 Hz, and 66.6 Hz have similar performance, with relative errors of 1.88%, 2.22%, and 2.88%, respectively. However, the maximum error among the 90 trials increases considerably between these frequencies, with maximum errors of 5.27%, 7.02%, and 8.25% for each respective frequency. Sampling frequencies of 50 Hz, 40 Hz, and 33.3 Hz also show good performance regarding the relative error, which is maintained below 5% for the 3 cases, but the maximum relative error found in these 3 frequencies is considerably higher than the found in the previous set.

In Table 4, MRE and maximum and minimum relative errors found among trials for the different sampling frequencies are shown. With slower sampling frequencies, MRE increases exponentially as shown in Figure 13, which suggests that sampling frequencies equal to or below 28.5 Hz are not reliable enough to maintain MRE below 5%. Also, sampling frequencies

slower than 50 Hz and 33.3 Hz show maximum relative error among trials higher than 10% and 20%, respectively.

Table 5 shows how MRE is distributed in different ranges. Only 200 Hz and 100 Hz sampling frequencies are able to maintain 95% of their results within 5% of relative error. Also, sampling

frequencies slower than 50 Hz considerably increase the percentage of relative errors found above 5%. These results suggest that sampling frequencies of 200 Hz and 100 Hz are the most reliable, frequencies of 66.6 Hz and 50 Hz have an acceptable performance, and the remaining sampling frequencies are unreliable for this specific application.

Figure 13. Mean relative error obtained for each proposed sampling frequency. As sampling frequency decreases, relative error increases exponentially.

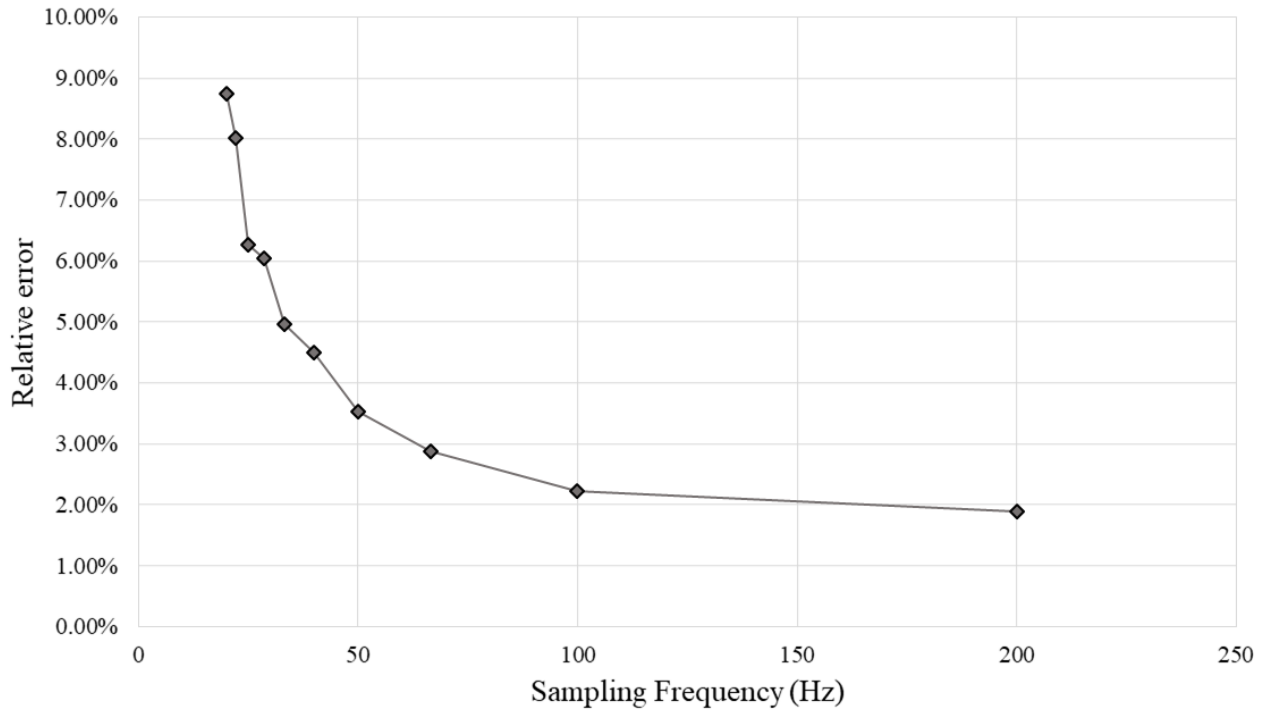


Table 4. Mean relative error and maximum and minimum relative errors obtained from all 90 trials for each sampling frequency analyzed.

Relative error	Sampling periods/frequencies									
	5 ms, 200 Hz	10 ms, 100 Hz	15 ms, 66.6 Hz	20 ms, 50 Hz	25 ms, 40 Hz	30 ms, 33.3 Hz	35 ms, 28.5 Hz	40 ms, 25 Hz	45 ms, 22.2 Hz	50 ms, 20 Hz
MRE <sup>a</sup>	1.88	2.22	2.88	3.52	4.50	4.97	6.04	6.27	8.02	8.75
MAX <sup>b</sup>	5.27	7.02	8.25	9.73	14.25	14.25	21.30	19.73	28.39	32.11
MIN <sup>c</sup>	0	0	0	0	0.70	0	0	0	0	0.73

<sup>a</sup>MRE: mean relative error.

<sup>b</sup>MAX: maximum relative error.

<sup>c</sup>MIN: minimum relative error.

Table 5. Percentage of trials whose relative error is within the ranges of 5% or less, higher than 5% and lower than 15%, and higher than 15%.

Relative error	Sampling periods/frequencies									
	5 ms, 200 Hz	10 ms, 100 Hz	15 ms, 66.6 Hz	20 ms, 50 Hz	25 ms, 40 Hz	30 ms, 33.3 Hz	35 ms, 28.5 Hz	40 ms, 25 Hz	45 ms, 22.2 Hz	50 ms, 20 Hz
RE <sup>a</sup> ≤5%	98.89	94.44	87.78	76.67	58.89	57.78	48.89	52.22	35.56	27.78
RE 5% to 15%	1.11	5.56	12.22	23.33	41.11	42.22	47.78	43.33	54.44	57.78
RE >15%	0	0	0	0	0	0	3.33	4.44	10	14.44

<sup>a</sup>RE: relative error.

When MRE is obtained from different jump heights ( $\leq 10$  cm, 10 to 20 cm, 20 to 30 cm, and  $>30$  cm) at each sampling frequency, the relevance of a proper sampling frequency when calculating height reached for small jumps ( $<20$  cm) is observed. When using sampling frequencies slower than 50 Hz, MRE obtained from these small jumps is always higher than 5%, and

for the slowest sampling frequency, MRE reaches a value of 21.50% for jump heights smaller than 10 cm. For higher jump heights, an increase in MRE ( $>5\%$ ) is noticeable for sampling frequencies slower than 40 Hz, reaching a value of up to 9.15% for the slowest sampling frequency. A summary of these results is shown in [Table 6](#).

**Table 6.** Mean relative error obtained from each of the analyzed sampling frequencies for different ranges of height reached during the vertical jump.

Jump height	Sampling periods/frequencies									
	5 ms, 200 Hz	10 ms, 100 Hz	15 ms, 66.6 Hz	20 ms, 50 Hz	25 ms, 40 Hz	30 ms, 33.3 Hz	35 ms, 28.5 Hz	40 ms, 25 Hz	45 ms, 22.2 Hz	50 ms, 20 Hz
<10 cm	1.31	2.29	4.33	2.62	5.28	5.09	5.40	6.29	10.77	21.50
10-20 cm	2.23	2.40	2.89	4.29	5.21	5.45	7.55	8.60	9.65	8.42
20-30 cm	1.91	2.31	3.02	3.55	4.62	5.29	5.06	6.12	6.96	9.15
>30 cm	1.60	1.95	2.50	2.92	3.63	4.17	5.83	4.34	7.32	6.85

## Discussion

### Principal Findings

The vertical jump is a test commonly used by health care professionals to assess strength in the lower limb muscles of a subject. Although this test is widely used for strength assessment among athletes, relevant information can be obtained from people with no relevant sports background.

An important point to highlight about the system developed in this study is its low price. The total for components used in construction is approximately US \$40. In comparison with commercially available devices, this developed system is

significantly more affordable. Among the devices commonly used on medical and sports fields to measure vertical jump height are the vertical jump test mat (Gill Athletics) [31], Just Jump system (Perform Better) [10,32], Vertec device (Gill Athletics) [10,33], electronic vertical jump tester (Gill Athletics) [34], Optojump testing (Perform Better) [12,35], and bilateral force plate (Hawkin Dynamics) [18,36]. Our proposed system will have to pass through different standards and certifications (like ISO standards [37]) before it can be considered as a standard medical device. [Table 7](#) shows a comparison of prices between commercially available devices and the system proposed here. Prices of the commercially available devices are listed as found at the moment of writing this article.

**Table 7.** Comparison of prices between commercially available devices and the system developed in this study.

Device for vertical jump measurement	Price \$
Proposal from this study (estimated price of components)	40
Vertical jump test mat (Gill Athletics) [31]	360
Just Jump System (Perform Better) [32]	629
Vertec device (Gill Athletics) [33]	760
Electronic vertical jump tester (Gill Athletics) [34]	2925
Optojump testing (Perform Better) [35]	3804
Bilateral force plate (Hawkin Dynamics) [36]	5000

Throughout data capturing in both experiments, some important points can be highlighted. Despite the advantages that the proposed system and reference device offer, both have an inherent error due to their sampling frequency (more specifically due to their sampling period). The proposed system has a sampling frequency of 200 Hz, and thus the sampling period is 5 ms. Likewise, the reference device has a sampling frequency of 120 Hz and a sampling period of 8.3 ms. This means that every sampling period each device updates its readings, which implies an uncertainty of the sampling period between data updates. In other words, there is an inherent uncertainty in the system during the takeoff and landing phases of the jump, time span that is used to calculate height reached. From both phases, the proposed system has a total uncertainty of 10 ms, while the

reference device has a total uncertainty of 16.6 ms. This inherent error is characteristic of electronic devices and directly related to their sampling frequency. Nonelectronic methods for jump height measuring lack of this inherent error, but as stated before, these methods tend to overestimate obtained measurements and are less precise.

Regarding the high-speed camera used as a reference device, when the recorded videos were analyzed, the ease of selecting the correct frames depended on the correct technique execution of the volunteer: taking off from both forefeet at the same time during the takeoff phase and landing with both forefeet at the same time during the landing phase. This was the ideal technique execution. On the other hand, some volunteers either took off or landed with only one forefoot and not with the same foot in

some occasions. In such cases, it was harder to select the takeoff and landing frames. This is difficult for volunteers to control without long-term training in the proper technique.

On the proposed protocol, the inclusion of two different jump techniques proved to be useful in order to increase the dynamic range of the data. The difference between the CMJ and CMJAS was significant. The addition of an arm swing increased jump height an average of 44.84% in the first experiment and 34.86% in the second experiment.

### Limitations

One of the main limitations of the developed system was its sampling frequency. Although the microcontroller used had a high CPU frequency, the sampling frequency was limited because of the number of operations needed for a full scan of the mat (16 rows and 16 columns, a scan of 256 individual cells) and number of operations this implies. Another limitation was the total sensing area of the system of 30×30 cm. Although no volunteer reported discomfort, the total area limits the stance of volunteers; in addition, the landing phase of every jump trial must be performed in a controlled manner, so the volunteer lands inside this area.

These limitations can be solved in future versions of the mat. The design of the mat can be modified to increase its total sensing area and the size of each row and column, so in this way with fewer number of rows and columns the same sensing area could be achieved, thus increasing the sampling frequency of the system. However, this would diminish the resolution of the system.

### Conclusions

In this study, a novel low-cost system for measurement of the jump height is proposed. Two experiments were

performed—one to validate the system and the other to assess the effects of different sampling frequencies.

When evaluating the performance of the proposed system in the first experiment, results show that with the proposed sampling frequency of 200 Hz relative error for all of the 228 jump trials is maintained below 5%. In the second experiment, with sampling frequencies of 200 Hz and 100 Hz, relative error is maintained below 5% for 98.89% and 94.44% of the jump trials, respectively.

The flight-time formula is a widely used, validated method to calculate height reached during vertical jumps. A high-speed camera as reference device has been used in related studies along with the flight-time formula, proving to be a reliable tool. Our first experiment showed through correlation and Bland-Altman analyses that the proposed system and a high-speed camera reference device produced systematically similar results when calculating jump height.

Our second experiment concluded that 200 Hz and 100 Hz sampling frequencies have similar performance, and both frequencies are reliable when calculating jump height using the flight-time formula. This implies that if access to hardware capable of processing data at 200 Hz were limited, hardware capable of processing data to at least 100 Hz could offer similar results. However, if higher sampling frequencies are available, they should be used.

These results demonstrate that the proposed system is as reliable as a commercially available device, and the selected sampling frequency of 200 Hz is reliable for obtaining relative errors below 5% for at least 95% of the jump trials. The proposed system offers an alternative for health care professionals to use a mobile monitoring station of their choice, and its price is more affordable than commercially available devices.

### Acknowledgments

This research was funded by grant Programa Operativo FEDER Construyendo Europa desde Aragón T49\_20R from the European Union and Gobierno de Aragón, grant UZCUD2019-TEC-02 from the Universidad de Zaragoza and Centro Universitario de la Defensa de Zaragoza, and grant 709365 from the Consejo Nacional de Ciencia y Tecnología.

### Conflicts of Interest

None declared.

### References

1. Rodrigues C, Correia M, Abrantes J, Nadal J, Benedetti Rodrigues MA. Consistency of surface electromyography assessment at lower limb selected muscles during vertical countermovement. *Annu Int Conf IEEE Eng Med Biol Soc* 2017 Jul;2017:402-405. [doi: [10.1109/EMBC.2017.8036847](https://doi.org/10.1109/EMBC.2017.8036847)] [Medline: [29059895](https://pubmed.ncbi.nlm.nih.gov/29059895/)]
2. Garcia M, Guzmán L, Valencia J, Henao V. Portable measurement system of vertical jump using an inertial measurement unit and pressure sensors. 2016 Presented at: 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA); 2016; Bucaramanga p. 1-5. [doi: [10.1109/stsiva.2016.7743299](https://doi.org/10.1109/stsiva.2016.7743299)]
3. Drazan J, Danielsen H, Vercelletto M, Loya A, Davis J, Eglash R. A case study for integrated STEM outreach in an urban setting using a do-it-yourself vertical jump measurement platform. *Annu Int Conf IEEE Eng Med Biol Soc* 2016 Aug;2016:3027-3030. [doi: [10.1109/EMBC.2016.7591367](https://doi.org/10.1109/EMBC.2016.7591367)] [Medline: [28268950](https://pubmed.ncbi.nlm.nih.gov/28268950/)]
4. Casartelli N, Müller R, Maffiuletti N. Validity and reliability of the Myotest accelerometric system for the assessment of vertical jump height. *J Strength Cond Res* 2010 Nov;24(11):3186-3193. [doi: [10.1519/JSC.0b013e3181d8595c](https://doi.org/10.1519/JSC.0b013e3181d8595c)] [Medline: [20940642](https://pubmed.ncbi.nlm.nih.gov/20940642/)]

5. Mijailovic N, Radakovic R, Peulic A, Milankovic I, Filipovic N. Using force plate, computer simulation and image alignment in jumping analysis. 2015 Presented at: 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE); 2015; Belgrade p. 1-4. [doi: [10.1109/bibe.2015.7367672](https://doi.org/10.1109/bibe.2015.7367672)]
6. Yahya U, Senanayake S, Naim A. Intelligent integrated wearable sensing mechanism for vertical jump height prediction in female netball players. 2017 Presented at: 2017 Eleventh International Conference on Sensing Technology (ICST); 2017; Sydney p. 1-7. [doi: [10.1109/icsenst.2017.8304483](https://doi.org/10.1109/icsenst.2017.8304483)]
7. García-Ramos A, Štirn I, Padiál P, Argüelles-Cienfuegos J, De la Fuente B, Strojnik V, et al. Predicting vertical jump height from bar velocity. *J Sports Sci Med* 2015 Jun;14(2):256-262 [FREE Full text] [Medline: [25983572](https://pubmed.ncbi.nlm.nih.gov/25983572/)]
8. Howard R, Conway R, Harrison A. Estimation of force during vertical jumps using body fixed accelerometers. 2014 Presented at: 25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CICT 2014); 2014; Limerick p. 102-107. [doi: [10.1049/cp.2014.0667](https://doi.org/10.1049/cp.2014.0667)]
9. Whitmer T, Fry A, Forsythe C, Andre M, Lane M, Hudy A, et al. Accuracy of a vertical jump contact mat for determining jump height and flight time. *J Strength Cond Res* 2015 Apr;29(4):877-881. [doi: [10.1519/JSC.0000000000000542](https://doi.org/10.1519/JSC.0000000000000542)] [Medline: [24852256](https://pubmed.ncbi.nlm.nih.gov/24852256/)]
10. Nuzzo J, Anning J, Scharfenberg J. The reliability of three devices used for measuring vertical jump height. *J Strength Cond Res* 2011 Sep;25(9):2580-2590. [doi: [10.1519/JSC.0b013e3181fee650](https://doi.org/10.1519/JSC.0b013e3181fee650)] [Medline: [21804426](https://pubmed.ncbi.nlm.nih.gov/21804426/)]
11. Balsalobre-Fernández C, Glaister M, Lockett RA. The validity and reliability of an iPhone app for measuring vertical jump performance. *J Sports Sci* 2015;33(15):1574-1579. [doi: [10.1080/02640414.2014.996184](https://doi.org/10.1080/02640414.2014.996184)] [Medline: [25555023](https://pubmed.ncbi.nlm.nih.gov/25555023/)]
12. Glatthorn J, Gouge S, Nussbaumer S, Stauffacher S, Impellizzeri F, Maffiuletti N. Validity and reliability of Optojump photoelectric cells for estimating vertical jump height. *J Strength Cond Res* 2011 Feb;25(2):556-560. [doi: [10.1519/JSC.0b013e3181ccb18d](https://doi.org/10.1519/JSC.0b013e3181ccb18d)] [Medline: [20647944](https://pubmed.ncbi.nlm.nih.gov/20647944/)]
13. Bui HT, Farinas M, Fortin A, Comtois A, Leone M. Comparison and analysis of three different methods to evaluate vertical jump height. *Clin Physiol Funct Imaging* 2015 May;35(3):203-209. [doi: [10.1111/cpf.12148](https://doi.org/10.1111/cpf.12148)] [Medline: [24690449](https://pubmed.ncbi.nlm.nih.gov/24690449/)]
14. Dias J, Dal Pupo J, Reis D, Borges L, Santos S, Moro A, et al. Validity of two methods for estimation of vertical jump height. *J Strength Cond Res* 2011 Jul;25(7):2034-2039. [doi: [10.1519/JSC.0b013e3181e73f6e](https://doi.org/10.1519/JSC.0b013e3181e73f6e)] [Medline: [21701288](https://pubmed.ncbi.nlm.nih.gov/21701288/)]
15. Kibele A. Possibilities and limitations in the biomechanical analysis of countermovement jumps: a methodological study. *J Appl Biomech* 1998;14(1):105-117. [doi: [10.1123/jab.14.1.105](https://doi.org/10.1123/jab.14.1.105)]
16. Boukhenous S, Attari M. A vertical jumping performance with and without arms swing by using a dynamometric platform. *Int Workshop Syst Signal Processing Appl* 2011:17-20. [doi: [10.1109/wosspa.2011.5931450](https://doi.org/10.1109/wosspa.2011.5931450)]
17. Zihajehzadeh S, Lee TJ, Lee JK, Hoskinson R, Park EJ. Integration of MEMS inertial and pressure sensors for vertical trajectory determination. *IEEE Trans Instrum Meas* 2015 Mar;64(3):804-814. [doi: [10.1109/tim.2014.2359813](https://doi.org/10.1109/tim.2014.2359813)]
18. Moir GL. Three different methods of calculating vertical jump height from force platform data in men and women. *Meas Phys Educ Exer Sci* 2008 Oct 15;12(4):207-218. [doi: [10.1080/10913670802349766](https://doi.org/10.1080/10913670802349766)]
19. Aragón LF. Evaluation of four vertical jump tests: methodology, reliability, validity, and accuracy. *Meas Phys Educ Exer Sci* 2000 Dec;4(4):215-228. [doi: [10.1207/s15327841mpee0404\\_2](https://doi.org/10.1207/s15327841mpee0404_2)]
20. Kassim A, Miskon M, Rahim N, Yasuno T. Effectiveness of reference height control system for tripod hopping robot. 2011 Presented at: 2011 4th International Conference on Mechatronics (ICOM); 2011; Kuala Lumpur p. 1-4. [doi: [10.1109/icom.2011.5937123](https://doi.org/10.1109/icom.2011.5937123)]
21. Leard J, Cirillo M, Katsnelson E, Kimiatek D, Miller T, Trebincevic K, et al. Validity of two alternative systems for measuring vertical jump height. *J Strength Cond Res* 2007 Nov;21(4):1296-1269. [doi: [10.1519/R-21536.1](https://doi.org/10.1519/R-21536.1)] [Medline: [18076265](https://pubmed.ncbi.nlm.nih.gov/18076265/)]
22. Balsalobre-Fernández C, Tejero-González CM, del Campo-Vecino J, Bavaresco N. The concurrent validity and reliability of a low-cost, high-speed camera-based method for measuring the flight time of vertical jumps. *J Strength Cond Res* 2014 Feb;28(2):528-533. [doi: [10.1519/JSC.0b013e318299a52e](https://doi.org/10.1519/JSC.0b013e318299a52e)] [Medline: [23689339](https://pubmed.ncbi.nlm.nih.gov/23689339/)]
23. Valle-Lopera DA, Castraño-Franco AF, Gallego-Londoño J, Hernández-Valdivieso AM. Test and fabrication of piezoresistive sensors for contact pressure measurement. *Rev Fac Ing Univ Antioquia* 2017 Mar(82):47-52. [doi: [10.17533/udea.redin.n82a06](https://doi.org/10.17533/udea.redin.n82a06)]
24. Giovanelli D, Farella E. Force sensing resistor and evaluation of technology for wearable body pressure sensing. *J Sensors* 2016;2016:1-13. [doi: [10.1155/2016/9391850](https://doi.org/10.1155/2016/9391850)]
25. Medrano-Sanchez C, Igual-Catalan R, Rodriguez-Ontiveros VH, Plaza-Garcia I. Circuit analysis of matrix-like resistor networks for eliminating crosstalk in pressure sensitive mats. *IEEE Sensors J* 2019 Sep 15;19(18):8027-8036. [doi: [10.1109/jsen.2019.2918592](https://doi.org/10.1109/jsen.2019.2918592)]
26. Young W. Laboratory strength assessment of athletes. 1995. URL: <https://www.worldathletics.org/download/downloadnsa?filename=9daf7a4d-c52b-4956-8453-caa4b378606c.pdf&urlslug=laboratory-strength-assessment-of-athletes> [accessed 2021-03-05]
27. Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. *J Strength Cond Res* 2004 Aug;18(3):551-555. [doi: [10.1519/1533-4287\(2004\)18<551:RAFVOS>2.0.CO;2](https://doi.org/10.1519/1533-4287(2004)18<551:RAFVOS>2.0.CO;2)] [Medline: [15320660](https://pubmed.ncbi.nlm.nih.gov/15320660/)]

28. Cheng KB, Wang C, Chen H, Wu C, Chiu H. The mechanisms that enable arm motion to enhance vertical jump performance—a simulation study. *J Biomech* 2008;41(9):1847-1854. [doi: [10.1016/j.jbiomech.2008.04.004](https://doi.org/10.1016/j.jbiomech.2008.04.004)] [Medline: [18514208](https://pubmed.ncbi.nlm.nih.gov/18514208/)]
29. Feltner ME, Frascetti DJ, Crisp RJ. Upper extremity augmentation of lower extremity kinetics during countermovement vertical jumps. *J Sports Sci* 1999 Jun;17(6):449-466. [doi: [10.1080/026404199365768](https://doi.org/10.1080/026404199365768)] [Medline: [10404494](https://pubmed.ncbi.nlm.nih.gov/10404494/)]
30. Harman EA, Rosenstein MT, Frykman PN, Rosenstein RM. The effects of arms and countermovement on vertical jumping. *Med Sci Sports Exerc* 1990 Dec;22(6):825-833. [doi: [10.1249/00005768-199012000-00015](https://doi.org/10.1249/00005768-199012000-00015)] [Medline: [2287261](https://pubmed.ncbi.nlm.nih.gov/2287261/)]
31. Vertical jump test mat. Gill Athletics. URL: [https://www.gillporter.com/gill\\_store/catalog/product/view/id/12661/](https://www.gillporter.com/gill_store/catalog/product/view/id/12661/) [accessed 2021-03-05]
32. Just jump system. Perform Better. URL: [https://www.performbetter.com/Just-Jump-System\\_2](https://www.performbetter.com/Just-Jump-System_2) [accessed 2021-03-05]
33. Vertec jump measuring device. Gill Athletics. URL: [https://www.gillporter.com/gill\\_store/vertec-jump-measuring-device.html](https://www.gillporter.com/gill_store/vertec-jump-measuring-device.html) [accessed 2021-03-05]
34. Electronic vertical jump tester. Gill Athletics. URL: [https://www.gillporter.com/gill\\_store/catalog/product/view/id/12673/](https://www.gillporter.com/gill_store/catalog/product/view/id/12673/) [accessed 2021-03-05]
35. Optojump jump testing. Perform Better. URL: <https://performbetter.co.uk/product/optojump/> [accessed 2021-03-05]
36. Hawkin dynamics G3 bilateral force plate set. Hawkin Dynamics. URL: <https://hawkindynamics.shop/products/hawkin-dynamics-g3-bilateral-force-plate-set> [accessed 2021-03-05]
37. ISO 13485 – medical devices. International Organization for Standardisation. URL: <https://www.iso.org/iso-13485-medical-devices.html> [accessed 2021-03-03]

## Abbreviations

- CMJ:** countermovement jump  
**CMJAS:** countermovement jump with arm swing  
**FSR:** force-sensitive resistor  
**MAE:** mean absolute error  
**MRE:** mean relative error

*Edited by L Buis; submitted 21.01.21; peer-reviewed by O Aragón Banderas, J Seitz; comments to author 02.03.21; revised version received 10.03.21; accepted 19.03.21; published 09.04.21.*

*Please cite as:*

*Vanegas E, Salazar Y, Igual R, Plaza I*

*Force-Sensitive Mat for Vertical Jump Measurement to Assess Lower Limb Strength: Validity and Reliability Study*

*JMIR Mhealth Uhealth* 2021;9(4):e27336

URL: <https://mhealth.jmir.org/2021/4/e27336>

doi: [10.2196/27336](https://doi.org/10.2196/27336)

PMID: [33835040](https://pubmed.ncbi.nlm.nih.gov/33835040/)

©Erik Vanegas, Yolocuauhtli Salazar, Raúl Igual, Inmaculada Plaza. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 09.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Relationship Between Major Depression Symptom Severity and Sleep Collected Using a Wristband Wearable Device: Multicenter Longitudinal Observational Study

Yuezhou Zhang<sup>1</sup>, MSc; Amos A Folarin<sup>1,2,3</sup>, PhD; Shaoxiong Sun<sup>1</sup>, PhD; Nicholas Cummins<sup>1</sup>, PhD; Rebecca Bendayan<sup>1,3</sup>, PhD; Yatharth Ranjan<sup>1</sup>, MSc; Zulqarnain Rashid<sup>1</sup>, PhD; Pauline Conde<sup>1</sup>, BSc; Callum Stewart<sup>1</sup>, MSc; Petroula Laiou<sup>1</sup>, PhD; Faith Matcham<sup>4</sup>, PhD; Katie M White<sup>4</sup>, BSc; Femke Lamers<sup>5</sup>, PhD; Sara Siddi<sup>6,7,8</sup>, PhD; Sara Simblett<sup>9</sup>, PhD, DCLinPsy; Inez Myin-Germeys<sup>10</sup>, PhD; Aki Rintala<sup>10,11</sup>, MSc; Til Wykes<sup>3,9</sup>, PhD; Josep Maria Haro<sup>6,7,8</sup>, MD; Brenda WJH Penninx<sup>5</sup>, PhD; Vaibhav A Narayan<sup>12</sup>, PhD; Matthew Hotopf<sup>3,4</sup>, PhD; Richard JB Dobson<sup>1,2,3</sup>, PhD; RADAR-CNS Consortium<sup>13</sup>

<sup>1</sup>Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

<sup>2</sup>Institute of Health Informatics, University College London, London, United Kingdom

<sup>3</sup>South London and Maudsley National Health Services Foundation Trust, London, United Kingdom

<sup>4</sup>Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

<sup>5</sup>Department of Psychiatry, Amsterdam Public Health Research Institute and Amsterdam Neuroscience, Amsterdam University Medical Centre, Vrije Universiteit and GGZ inGeest, Amsterdam, Netherlands

<sup>6</sup>Teaching Research and Innovation Unit, Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain

<sup>7</sup>Centro de Investigación Biomédica en Red de Salud Mental, Madrid, Spain

<sup>8</sup>Faculty of Medicine and Health Sciences, Universitat de Barcelona, Barcelona, Spain

<sup>9</sup>Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

<sup>10</sup>Center for Contextual Psychiatry, Department of Neurosciences, Katholieke Universiteit Leuven, Leuven, Belgium

<sup>11</sup>Faculty of Social Services and Health Care, LAB University of Applied Sciences, Lahti, Finland

<sup>12</sup>Janssen Research and Development LLC, Titusville, NJ, United States

<sup>13</sup>see Acknowledgments

## Corresponding Author:

Richard JB Dobson, PhD

Department of Biostatistics & Health Informatics

Institute of Psychiatry, Psychology and Neuroscience

King's College London

SGDP Centre, IoPPN, Box PO 80

De Crespigny Park, Denmark Hill

London

United Kingdom

Phone: 44 20 7848 0473

Email: [richard.j.dobson@kcl.ac.uk](mailto:richard.j.dobson@kcl.ac.uk)

## Abstract

**Background:** Sleep problems tend to vary according to the course of the disorder in individuals with mental health problems. Research in mental health has associated sleep pathologies with depression. However, the gold standard for sleep assessment, polysomnography (PSG), is not suitable for long-term, continuous monitoring of daily sleep, and methods such as sleep diaries rely on subjective recall, which is qualitative and inaccurate. Wearable devices, on the other hand, provide a low-cost and convenient means to monitor sleep in home settings.

**Objective:** The main aim of this study was to devise and extract sleep features from data collected using a wearable device and analyze their associations with depressive symptom severity and sleep quality as measured by the self-assessed Patient Health Questionnaire 8-item (PHQ-8).

**Methods:** Daily sleep data were collected passively by Fitbit wristband devices, and depressive symptom severity was self-reported every 2 weeks by the PHQ-8. The data used in this paper included 2812 PHQ-8 records from 368 participants recruited from 3 study sites in the Netherlands, Spain, and the United Kingdom. We extracted 18 sleep features from Fitbit data that describe participant sleep in the following 5 aspects: sleep architecture, sleep stability, sleep quality, insomnia, and hypersomnia. Linear mixed regression models were used to explore associations between sleep features and depressive symptom severity. The  $z$  score was used to evaluate the significance of the coefficient of each feature.

**Results:** We tested our models on the entire dataset and separately on the data of 3 different study sites. We identified 14 sleep features that were significantly ( $P < .05$ ) associated with the PHQ-8 score on the entire dataset, among them awake time percentage ( $z = 5.45$ ,  $P < .001$ ), awakening times ( $z = 5.53$ ,  $P < .001$ ), insomnia ( $z = 4.55$ ,  $P < .001$ ), mean sleep offset time ( $z = 6.19$ ,  $P < .001$ ), and hypersomnia ( $z = 5.30$ ,  $P < .001$ ) were the top 5 features ranked by  $z$  score statistics. Associations between sleep features and PHQ-8 scores varied across different sites, possibly due to differences in the populations. We observed that many of our findings were consistent with previous studies, which used other measurements to assess sleep, such as PSG and sleep questionnaires.

**Conclusions:** We demonstrated that several derived sleep features extracted from consumer wearable devices show potential for the remote measurement of sleep as biomarkers of depression in real-world settings. These findings may provide the basis for the development of clinical tools to passively monitor disease state and trajectory, with minimal burden on the participant.

(*JMIR Mhealth Uhealth* 2021;9(4):e24604) doi:[10.2196/24604](https://doi.org/10.2196/24604)

## KEYWORDS

mobile health (mHealth); mental health; depression; sleep; wearable device; monitoring

## Introduction

According to the report of the World Health Organization, the total number of people with depression was estimated to exceed 300 million in 2015, equivalent to 4.4% of the world's population [1]. There are several depression-related adverse outcomes, including premature mortality [2], decline in quality of life [3], and loss of occupational function [4].

Sleep disturbances are prevalent among depression patients; more than 90% of patients with depression reported poor sleep quality [5]. Sleep disturbances cover a wide range of different symptoms and disorders including insomnia, hypersomnia, excessive daytime sleepiness, and circadian rhythm disturbance [6]. Insomnia and sleep quality have been observed to be bidirectionally related to depression in several longitudinal studies [6]. Hypersomnia is more frequently present in depressive episodes of bipolar patients [7,8]. Changes in sleep architecture, such as reduced deep sleep, increased rapid eye movement (REM) sleep, and shortened REM latency, are also significant predictors of depression [9,10].

The gold standard for sleep evaluation is polysomnography (PSG), which involves several physiological measurements including electroencephalogram, electrocardiogram, electromyogram, and accelerometers [11]. Using PSG to assess sleep lacks ecological validity and is time-consuming, expensive, and labor-intensive, requiring dedicated equipment and separate laboratory rooms as well as experts to analyze the physiological signals. Since depression can affect patients for an extended period, long-term monitoring of sleep quality is essential. Due to the above shortcomings, PSG is not suitable for long-term sleep monitoring [12]. A sleep questionnaire, such as the Pittsburgh Sleep Quality Index (PSQI) [13], is another useful method to assess sleep. This method relies on the self-reporting of subjective factors, like low recall of sleep, that may affect the accuracy of the assessment [14].

Several recent studies have used wearable devices to estimate sleep quality and sleep-related parameters [15-18] and analyzed the relationship between sleep and depression [19-21]. Miwa et al [19] estimated sleep quality by detecting rollover movements during sleep and observed a significant difference in sleep quality between nondepressed and depressed people. Mark et al [20] estimated the sleep duration of 40 information workers for 12 days using a Fitbit wristband and found that sleep duration was positively correlated with mood. DeMasi et al [21] found that sleep was significantly related to changes in depressive symptoms. These studies have mostly been performed on single center and relatively small datasets (number of participants fewer than 100). Moreover, most of these studies only used basic sleep parameters, such as sleep duration; detailed information on sleep architecture, sleep patterns, and stability of sleep was not considered. The relationship between detailed sleep features, as estimated from data supplied by wearable devices, and depression is yet to be fully explored.

The first aim of this study was to design more sleep-related features, from wearable device data, that reflect the sleep architecture, sleep stability, sleep quality, and sleep disturbances (insomnia and hypersomnia) of the participant. The second aim was to explore associations between these sleep features and depressive symptom severity on a relatively large, multisite dataset. The third aim was to compare our findings with previous studies that used other measurements to assess sleep such as PSG and sleep questionnaires.

## Methods

### Dataset

#### Study Participants and Settings

The data we used in this paper were collected from a major EU Innovative Medicines Initiative research project, Remote Assessment of Disease and Relapse-Central Nervous System (RADAR-CNS) [22]. This project aims to investigate the use

of remote measurement technologies to monitor people with depression, epilepsy, and multiple sclerosis in real-world settings. The study protocol for the depression component (Remote Assessment of Disease and Relapse–Major Depressive Disorder [RADAR-MDD]) is described in detail in Matcham et al [23]. The RADAR-MDD project aims to recruit 600 participants with a recent history of depression in 3 study sites (King’s College London [KCL], UK; Vrije Universiteit Medisch Centrum [VUMC], Amsterdam, The Netherlands; and Centro de Investigación Biomédica en Red [CIBER], Barcelona, Spain). Recruitment procedures vary slightly across sites and eligible participants are identified either through existing research cohorts (in KCL and VUMc) who had given consent to be contacted for research purposes; advertisements in general practices, psychologist practices, newspapers, and Hersenonderzoek.nl [24], which is a Dutch online registry (VUMc); or through mental health services (in KCL and CIBER) [23]. Participants from KCL and VUMc are community-based, while the participants from CIBER come from a clinical population. As part of the study, participants are asked to install several remote monitoring technology apps and use an activity tracker for up to 2 years of follow-up. Many categories of passive and active data are being collected and uploaded to an open-source platform, RADAR-base [25]. In this paper, we focus on the sleep and Patient Health Questionnaire 8-item (PHQ-8) data [26].

### **Sleep Data**

According to the American Academy of Sleep Medicine manual for the scoring of sleep and associated events, sleep can be divided into 2 phases, REM sleep and non-REM (NREM) sleep, and NREM sleep can be subdivided into N1, N2, and N3 stages according to characteristic patterns of brain waves collected by PSG [11]. In our project, the daily sleep records of participants were collected by the Charge 2 or Charge 3 (Fitbit Inc). An entire night’s sleep is divided into 4 stages: awake, light, deep, and REM. The light stage provides estimates for the N1 and N2 stages in PSG, while the deep stage provides estimates for the N3 stage in PSG. According to several validation studies of Fitbit, the Fitbit wristband had limited specificity in sleep stages estimation [27–29]. Therefore, in this study, we were not expecting the Fitbit devices to provide information as accurate as PSG would have provided. However, the Fitbit devices were deemed sensitive enough to detect changes in sleep-wake states [27–29]; therefore, the provided sleep stage information could be used to determine estimates for detailed sleep parameters based on known sleep pathology.

### **PHQ-8 Data**

The variability of each participant’s depressive symptom severity was measured via the PHQ-8, conducted by mobile phone every 2 weeks. The questionnaire contains 8 questions, with the score of each subitem ranging from 0 to 3. The total score (range 0 to 24) of all subitems is the PHQ-8 score, which can evaluate depressive symptom severity of the participant for the past 2 weeks. A PHQ-8 score  $\geq 10$  is the most commonly recommended cutpoint for clinically significant depressive symptoms [26] (ie, if the PHQ-8 of a participant is  $\geq 10$ , the participant is likely to have had depressive symptoms in the previous 2 weeks). In the PHQ-8, subitem 3 refers to sleep. The content of subitem 3 is “Trouble falling or staying asleep, or sleeping too much” [26]. A higher score in subitem 3 indicates worse self-reported sleep in the past 2 weeks. For reading convenience, we denoted the score of subitem 3 as the sleep subscore in this paper.

### **Sociodemographics**

Sociodemographic of participants were collected during the enrollment session. According to previous studies on the associations between depression and sociodemographic characteristics [30,31], we considered baseline age, gender, education level, and annual income as potential confounding variables in our analyses. Due to the different educational systems in different countries, we simply divided the education level into 2 levels: degree (or above) and below degree. The annual income levels of Spain and the Netherlands were transformed into equivalent British levels.

### **Feature Extraction**

#### **Feature Window Size**

For each PHQ-8 record, we extracted sleep features from a 2-week time window prior to the PHQ-8 completion time, as the PHQ-8 score is used to represent the depressive symptom severity of the participant for the past 2 weeks. The feature window is denoted as  $\Delta t$  in this paper.

#### **Sleep Features**

According to known sleep pathology and our experience, 18 sleep features extracted in this paper were divided into the following 5 categories (Table 1): sleep architecture, representing the basic and cyclical patterns of sleep; sleep stability, representing the variance of sleep in the feature window; sleep quality, measures relating to total sleep and wake times; insomnia, trouble falling or staying asleep; and hypersomnia, excessive sleepiness.

**Table 1.** A list of sleep features used in this study and their short descriptions.

Features	Description	Unit
<b>Sleep architecture</b>		
Av_tst	Mean total sleep time	Hour
Av_time_bed	Mean time in bed	Hour
Deep_pct	Mean percentage of deep sleep	%
Light_pct	Mean percentage of light sleep	%
REM_pct	Mean percentage of REM <sup>a</sup> sleep	%
NREM_pct	Mean percentage of NREM <sup>b</sup> sleep	%
Awake_pct	Mean percentage of awake time	%
Av_onset	Mean sleep onset time	Hour
Av_offset	Mean sleep offset time	Hour
REM_L	Mean REM latency time	Hour
<b>Sleep stability</b>		
Std_tst	Standard deviation of total sleep time	Hour
Std_onset	Standard deviation of sleep onset time	Hour
Std_offset	Standard deviation of sleep offset time	Hour
<b>Sleep quality</b>		
Efficiency	Mean sleep efficiency	%
Awake_5	Mean number of awakenings (>5 minutes) per night	Times
WKD_diff	Total sleep time difference between weekend and weekdays	Hour
<b>Insomnia</b>		
M_insomnia	Percentage of days with potential middle insomnia	%
<b>Hypersomnia</b>		
Dur_10	Percentage of days with total sleep time >10 hours	%

<sup>a</sup>REM: rapid eye movement.

<sup>b</sup>Non-REM: non-rapid eye movement.

### Sleep Architecture

The features of sleep architecture were intended to describe the basic and cyclical patterns of sleep. Therefore, we extracted some features similar to those in the PSG report (total sleep time, time in bed, sleep onset time, sleep offset time, and REM latency) [32], and features of the percentages of all sleep stages. Total sleep time of one night is defined as the sum of all nonawake stages (light, deep, and REM) [32]. The mean total sleep time in  $\Delta t$  was denoted as *Av\_tst*. Time in bed of one night is defined as the sum of all sleep stages (awake, light, deep, and REM) of the entire night [32]. The mean time in bed in  $\Delta t$  was denoted as *Av\_time\_bed*. Percentage of each sleep stage is defined as the percentage of the time in the sleep stage to the time in bed of the entire night. Different sleep stages have different functions and can reflect the quality of sleep. Deep sleep is considered essential for memory consolidation [33], and REM sleep favors the preservation of memory [34]. A previous sleep report has shown that more deep sleep and fewer awakenings represent better sleep quality [32]. Therefore, we extracted the mean percentages of these 4 sleep stages in  $\Delta t$ , and denoted them as *Deep\_pct*, *Light\_pct*, *REM\_pct*, *Awake\_pct*,

respectively. The combination of deep and light sleep is NREM sleep. The mental activity that occurs in NREM and REM sleep is a result of 2 different mind generators, which also explains the difference in mental activity [35]. So, we extracted the mean percentage of NREM sleep in  $\Delta t$ , which was denoted as *NREM\_pct*. We calculated the mean sleep onset time (the first nonawake stage) in  $\Delta t$ , denoted as *Av\_onset*. Mean sleep offset time (the last nonawake stage) in  $\Delta t$  was calculated and denoted as *Av\_offset*. Previous literature has shown that shortened REM latency can be considered as a biological mark of depression relapse [9]. REM latency is defined as the interval between sleep onset and occurrence of the first REM stage. The mean REM latency in  $\Delta t$  was denoted as *REM\_L*.

### Sleep Stability

The features in this category were used to estimate the variance of sleep during  $\Delta t$ . We extracted the standard deviation of total sleep time, sleep onset time, and sleep offset time in  $\Delta t$ , which were denoted as *Std\_tst*, *Std\_onset*, and *Std\_offset*, respectively.

## Sleep Quality

In this paper, we used features of sleep efficiency, awakenings, and weekend catch-up sleep to describe sleep quality. The definition of sleep efficiency is the percentage of total sleep time to time in bed [32]. Mean sleep efficiency in  $\Delta t$  was denoted as *Efficiency*. The definition of awakenings (>5 minutes) for one night is the number of episodes in which an individual is awake for more than 5 minutes [32]. The average number of awakenings in  $\Delta t$  was denoted as *Awake\_5*. Weekend catch-up sleep is an indicator of insufficient weekday sleep, which might be associated with depression level [36]. A longer total sleep time during the weekend compared with weekdays may reflect the actual sleep needed [37]. Therefore, we calculated the mean total sleep time difference between weekend and weekdays in  $\Delta t$ , which was denoted as *WKD\_diff*.

## Insomnia

A review of several longitudinal studies suggested that insomnia is bidirectionally related to depression [6]. According to the diagnostic features provided in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* [38], insomnia manifests as initial insomnia (difficulty initiating sleep at bedtime), middle insomnia (frequent or prolonged awakening throughout the night), and late insomnia (early-morning awakening with an inability to return to sleep).

For initial insomnia and late insomnia, mean sleep onset time (*Av\_onset*) and sleep offset time (*Av\_offset*) can be used to partially reflect them, respectively. We define potential middle insomnia to be whether the total sleep time is less than 6 hours and there is at least one prolonged awakening ( $\geq 30$  minutes) during the night. The percentage of days with potential middle insomnia in the feature window was denoted as *M\_Insomnia*.

## Hypersomnia

Hypersomnia can be another symptom of depression [7]. The hypersomnia criteria used in Tam et al [39] is sleeping more than 10 hours per day, 3 days per week. In this paper, the percentage of days with total sleep time greater than 10 hours was extracted in  $\Delta t$  and denoted as *Dur\_10*.

## Statistical Method

### Data Inclusion Criteria

Sleep and PHQ-8 records were missing in our data cohort for a variety of expected reasons, including the participants not wearing the Fitbit wristband when they slept, participants forgetting to complete the PHQ-8, and the Fitbit wristband being damaged during follow-up. We, therefore, specified the following inclusion criteria: (1) PHQ-8 record should be completed (ie, participant answered all 8 questions in the questionnaire); (2) number of days with sleep records in the feature window should be at least 12 days (approximately 85% of the feature window size) [40]; (3) number of PHQ-8 records for each participant should be greater than or equal to 3 [41]; (4) date of PHQ-8 records should be before February 2020, because the impact of the COVID-19 pandemic on sleep needs to be excluded [42].

## Statistical Analyses

In our study, each participant had multiple PHQ-8 records and repeated sleep measures. For this reason, we used linear mixed models, which allow for accounting of both within and between-individual variability over time [43]. For each sleep feature, a 3-level linear mixed model with a participant-specific random intercept and a site-specific random intercept was built on the entire dataset to explore the association between this sleep feature and depressive symptom severity (PHQ-8) by bivariate analysis. We then used 2-level linear mixed models with participant-specific random intercepts to test these associations on the 3 subsets (KCL, CIBER, and VUmc) separately. We similarly analyzed the associations between sleep features and sleep subscore. All models were adjusted for baseline age, gender, education level, and annual income, which were specified as fixed effects. Model assumptions were checked by the histograms of residuals and Q-Q plots. If the residuals are not normally distributed, the Box-Cox transformation was performed [44]. The  $z$  score was used to evaluate the statistical significance of the coefficient of each model. All  $P$  values of these tests were corrected by using the Benjamini-Hochberg method [45] for multiple comparisons, and the significance level of the corrected  $P$  value was set to .05. All linear mixed models were implemented by using the lme4 package for R software version 3.6.1 (R Foundation for Statistical Computing).

In order to identify and compare the relationship between self-reported sleep and self-reported depression among different study sites, Spearman correlations were calculated between the PHQ-8 score and sleep subscore on the 3 study sites separately.

An example of such a 3-level linear mixed model is as follows:

$$\text{Sleep}_{ijk} = \delta_{000} + V_{00k} + U_{0jk} + \beta_1(\text{PHQ8}_{ijk}) + \beta_2(\text{age}_{jk}) + \beta_3(\text{gender}_{jk}) + \beta_4(\text{education}_{jk}) + \beta_5(\text{income}_{jk}) + \epsilon_{ijk}$$

where  $\text{PHQ8}_{ijk}$  is the  $i^{\text{th}}$  PHQ-8 score of the participant  $j$  of the site  $k$ ,  $\text{Sleep}_{ijk}$  is one sleep feature extracted in  $\Delta t$  before the  $i^{\text{th}}$  PHQ-8 record of the participant  $j$  of the site  $k$ ,  $\text{age}_{jk}$ ,  $\text{gender}_{jk}$ ,  $\text{education}_{jk}$ , and  $\text{income}_{jk}$  are potential confounding variables of the participant  $j$  of the site  $k$ ,  $\epsilon_{ijk}$  is the residual,  $\delta_{000}$  is the fixed effect on intercept,  $U_{0jk}$  is the random intercept of the participant  $j$  in the site  $k$ , and  $V_{00k}$  is the random intercept of the site  $k$ .

## Results

### Data Summary

According to our data inclusion criteria, from June 2018 to February 2020, 2812 PHQ-8 records from 368 participants collected from 3 study sites were included for our analysis. A summary of the sociodemographic characteristics of these participants at baseline and scores of all PHQ-8 records is shown in Table 2. The Kruskal-Wallis test was used to determine whether there were any significant differences for these characteristics between the sites. These tests revealed that, except for gender, sociodemographic characteristics and distribution of PHQ-8 scores differed between the study sites. The histograms of PHQ-8 scores of the study sites and the entire

dataset are shown in [Figure 1](#). We can observe that the KCL site had the most PHQ-8 records among the sites. PHQ-8 scores from the CIBER site were relatively high, probably because participants in the CIBER site came from a clinical population. [Figure 2](#) presents pairwise Spearman correlation coefficients between all 18 sleep features. [Table 3](#) shows the results of

Spearman correlation analysis; we can observe there was a strong positive correlation between the sleep subscore and PHQ-8 score ( $r=.73$ ,  $z=54.48$ ,  $P<.001$ ) on the entire dataset, but this correlation was relatively weaker on the VUmc data ( $r=.64$ ,  $z=18.75$ ,  $P<.001$ ).

**Table 2.** A summary of sociodemographic characteristics and PHQ-8 records of participants from the 3 study sites and results of Kruskal-Wallis tests on these characteristics.

Characteristic	KCL <sup>a</sup>	CIBER <sup>b</sup>	VUmc <sup>c</sup>	P value <sup>d</sup>
Participants, n	189	96	83	— <sup>e</sup>
PHQ-8 <sup>f</sup> records, n	1547	708	557	—
PHQ-8 scores, median (Q1, Q3)	8 (4, 12)	14 (8, 19)	9 (5, 13)	<.001
The PHQ-8 score $\geq 10$ , n (%)	599 (38.7)	492 (69.5)	248 (44.5)	<.001
Age at baseline, median (Q1, Q3)	46 (30.3, 59.0)	55 (49.3, 60.8)	42 (28.0, 57.0)	<.001
Female sex, n (%)	144 (76.2)	69 (71.9)	65 (81.9)	.62
<b>Education<sup>g</sup>, n (%)</b>	—	—	—	<.001
Degree or above	116 (61.4)	21 (21.9)	40 (48.2)	—
Below degree	73 (38.6)	75 (78.1)	43 (51.8)	—
<b>Annual income<sup>h</sup> (£), n (%)</b>	—	—	—	.009
<15,000	40 (21.2)	28 (29.2)	24 (28.9)	—
15,000-40,000	80 (42.3)	53 (55.2)	34 (41.0)	—
>40,000	67 (35.5)	15 (15.6)	14 (16.9)	—
Not mentioned	2 (1.1)	0 (0)	11 (13.3)	—

<sup>a</sup>KCL: King's College London.

<sup>b</sup>CIBER: Centro de Investigación Biomédica en Red.

<sup>c</sup>VUmc: Vrije Universiteit Medisch Centrum.

<sup>d</sup>P value of Kruskal-Wallis test.

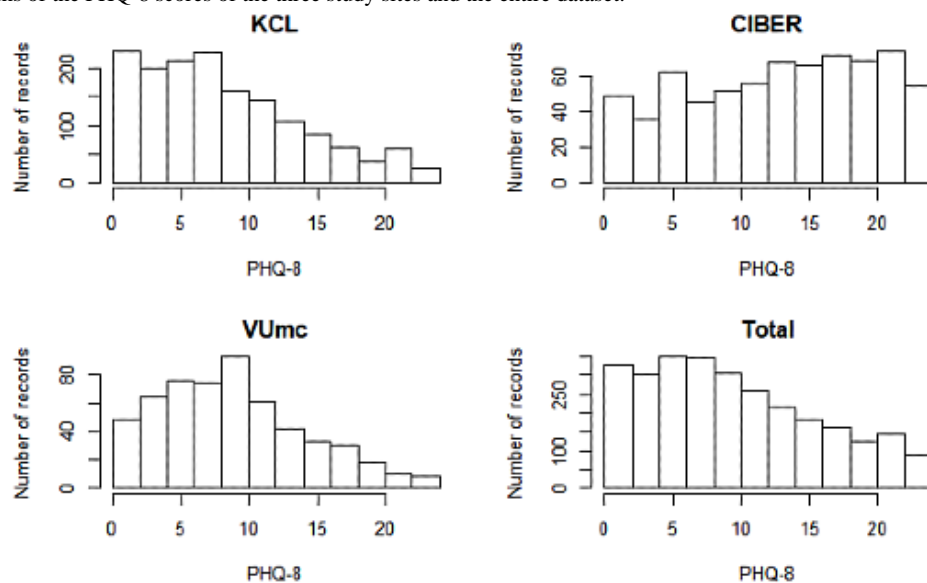
<sup>e</sup>Not applicable.

<sup>f</sup>PHQ-8: Patient Health Questionnaire 8-item.

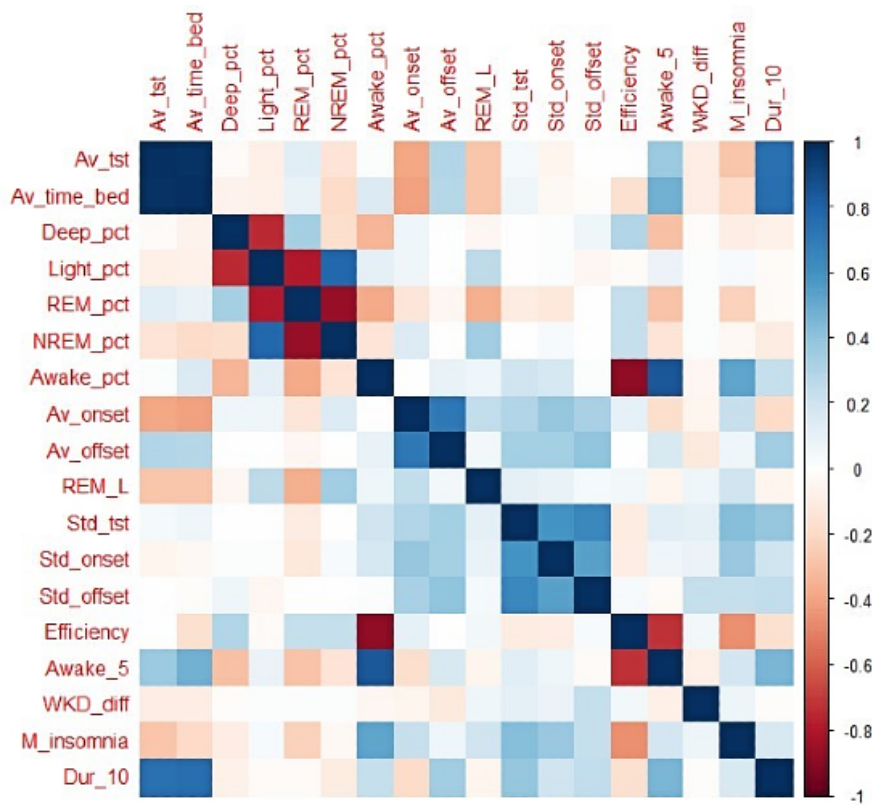
<sup>g</sup>Education levels of Spain and the Netherlands transformed into equivalent British education levels.

<sup>h</sup>Annual income levels of Spain and the Netherlands transformed into equivalent British levels.

**Figure 1.** Histograms of the PHQ-8 scores of the three study sites and the entire dataset.



**Figure 2.** Correlation plot of pairwise Spearman correlations between all sleep features. Descriptions of abbreviations of sleep features are shown in Table 1.



**Table 3.** Spearman correlation coefficients between the PHQ-8 score and sleep subscore<sup>a</sup> on the 3 study sites and their 95% confidence intervals, z score statistics, and P values.

Study site	r	95% CI	z score	P value
KCL <sup>b</sup>	.74	0.71, 0.76	41.99	<.001
CIBER <sup>c</sup>	.78	0.75, 0.81	32.09	<.001
VUmc <sup>d</sup>	.64	0.58, 0.69	18.75	<.001
Total	.73	0.71, 0.74	54.48	<.001

<sup>a</sup>Sleep subscore represents the score of subitem 3 in the PHQ-8.

<sup>b</sup>KCL: King’s College London.

<sup>c</sup>CIBER: Centro de Investigación Biomédica en Red.

<sup>d</sup>VUmc: Vrije Universiteit Medisch Centrum.

### Three-Level Linear Mixed Models on the Entire Dataset

Table 4 shows the results from 3-level linear mixed regression models that reflect the associations between sleep features and the PHQ-8 score and sleep subscore, respectively. A total of 14 sleep features were found to be significantly associated with the PHQ-8 score, among them awake percentage ( $z=5.45$ ,

$P<.001$ ), awakening times ( $z=5.53$ ,  $P<.001$ ), insomnia ( $z=4.55$ ,  $P<.001$ ), mean sleep offset time ( $z=6.19$ ,  $P<.001$ ), and hypersomnia ( $z=5.30$ ,  $P<.001$ ) were the top 5 features ranked by z score statistics. The percentages of light sleep (*Light\_pct*) and NREM sleep (*NREM\_pct*) and sleep efficiency (*Efficiency*) were significantly and negatively associated with the PHQ-8 score, whereas the rest of the significant features were positively associated with the PHQ-8 score.

**Table 4.** Slope coefficient estimates, 95% confidence intervals, z score statistics, and *P* values from 3-level linear mixed models on the entire dataset for exploring associations between sleep features<sup>a</sup> and the PHQ-8 score and sleep subscore<sup>b</sup>.

Features	PHQ-8 <sup>c</sup> score				Sleep subscore			
	Coeff. <sup>d</sup>	95% CI	z score	<i>P</i> value	Coeff.	95% CI	z score	<i>P</i> value
Av_tst	0.013	0.006, 0.019	3.93	<.001	-0.004	-0.034, 0.025	-0.28	.78
Av_time_bed	0.016	0.009, 0.023	4.45	<.001	0.005	-0.028, 0.038	0.29	.77
Deep_pct	-0.007	-0.026, 0.011	-0.75	.45	-0.104	-0.191, -0.017	-2.34	.02
Light_pct	-0.032	-0.064, -0.001	-2.02	.04	0.090	-0.057, 0.237	1.20	.23
REM_pct	0.003	-0.021, 0.027	0.25	.80	-0.125	-0.238, -0.012	-2.17	.03
NREM_pct	-0.038	-0.062, -0.014	-3.12	.002	-0.014	-0.127, 0.098	-0.25	.80
Awake_pct	0.035	0.022, 0.048	5.45	<.001	0.139	0.079, 0.199	4.58	<.001
Av_onset	0.007	-0.001, 0.015	1.71	.09	0.078	0.040, 0.115	4.03	<.001
Av_offset	0.025	0.017, 0.033	6.19	<.001	0.097	0.060, 0.135	5.10	<.001
REM_L	0.034	-0.021, 0.088	1.21	.23	0.085	-0.178, 0.347	0.63	.53
Std_tst	0.008	0.004, 0.012	4.07	<.001	0.047	0.028, 0.067	4.77	<.001
Std_onset	0.012	0.004, 0.019	3.11	.002	0.060	0.022, 0.097	3.13	.002
Std_offset	0.012	0.005, 0.018	3.58	<.001	0.069	0.037, 0.100	4.26	<.001
Efficiency	-0.025	-0.037, -0.012	-3.91	<.001	-0.108	-0.167, -0.050	-3.65	<.001
Awake_5	0.016	0.010, 0.022	5.53	<.001	0.038	0.011, 0.065	2.77	.006
WKD_diff	0.134	0.039, 0.230	2.76	.006	0.747	0.255, 1.240	2.98	.003
M_insomnia	0.370	0.211, 0.530	4.55	<.001	2.373	1.595, 3.151	5.98	<.001
Dur_10	0.309	0.195, 0.423	5.30	<.001	0.909	0.357, 1.462	3.23	.001

<sup>a</sup>Definitions of sleep features in this table are shown in Table 1.

<sup>b</sup>Sleep subscore represents the score of subitem 3 in the PHQ-8.

<sup>c</sup>PHQ-8: Patient Health Questionnaire 8-item.

<sup>d</sup>Slope coefficient estimates for all sleep features.

For sleep subscore, we can notice that deep sleep percentage (*Deep\_pct*), REM sleep percentage (*REM\_pct*), and sleep efficiency (*Efficiency*) were significantly and negatively associated with the sleep subscore, whereas features of the percentage of awake time (*Awake\_pct*), unstable sleep (*Std\_tst*, *Std\_onset*, *Std\_offset*), awakening times (*Awake\_5*), weekend catch-up sleep (*WKD\_diff*), sleep onset time (*Av\_onset*), sleep offset time (*Av\_offset*), insomnia (*M\_insomnia*), and hypersomnia (*Dur\_10*) were significantly and positively associated with the sleep subscore.

## Two-Level Linear Mixed Models on Different Research Sites

Table 5 provides the results from 2-level linear mixed models which show the associations between sleep features and the

PHQ-8 score on different research sites separately. On the KCL data, most associations between sleep features and depression were consistent with the results on the entire dataset. On the CIBER data, some features were no longer significantly associated with the PHQ-8 score. However, on the VUmc data, most features lost their significance except features of total sleep time (*Av\_tst*), time in bed (*Av\_time\_bed*), REM latency (*REM\_L*), and awakenings (*Awake\_5*).

Table 6 shows associations between sleep features and the sleep subscore on different research sites. The significance of associations between sleep features and the sleep subscore were different among the 3 study sites. Notably, the insomnia feature (*M\_insomnia*) and at least one feature of sleep stability were significantly positively associated with sleep subscore on the data of all 3 sites.



**Table 5.** Coefficient estimates, 95% confidence intervals, and *P* values from 2-level linear mixed models on the 3 study sites for exploring associations between sleep features<sup>a</sup> and the PHQ-8 score.

Features	KCL <sup>b</sup>			CIBER <sup>c</sup>			VUmc <sup>d</sup>		
	Coeff. <sup>e</sup>	95% CI	<i>P</i> value	Coeff.	95% CI	<i>P</i> value	Coeff.	95% CI	<i>P</i> value
Av_tst	0.013	0.005, 0.020	.001	0.016	-0.001, 0.033	.06	0.011	0, 0.022	.049
Av_time_bed	0.016	0.008, 0.024	<.001	0.021	0.002, 0.040	.03	0.013	0.001, 0.025	.04
Deep_pct	-0.005	-0.028, 0.018	.69	0.024	-0.022, 0.071	.31	-0.037	-0.074, 0.001	.06
Light_pct	-0.046	-0.087, -0.006	.03	-0.081	-0.155, -0.007	.03	0.019	-0.043, 0.082	.55
REM_pct	0.013	-0.018, 0.043	.43	0.015	-0.042, 0.071	.62	-0.007	-0.055, 0.041	.77
NREM_pct	-0.049	-0.080, -0.018	.002	-0.060	-0.116, -0.005	.04	-0.016	-0.062, 0.030	.50
Awake_pct	0.037	0.020, 0.054	<.001	0.043	0.015, 0.071	.003	0.022	-0.003, 0.047	.09
Av_onset	0.010	0.000, 0.020	.047	0.004	-0.018, 0.025	.74	-0.005	-0.021, 0.010	.52
Av_offset	0.029	0.018, 0.039	<.001	0.024	0.004, 0.043	.02	0.012	-0.004, 0.029	.14
REM_L	0.019	-0.049, 0.088	.58	0.106	-0.026, 0.237	.12	-0.126	-0.231, -0.020	.02
Std_tst	0.008	0.003, 0.013	.001	0.009	0, 0.019	.06	0.002	-0.006, 0.010	.62
Std_onset	0.007	-0.002, 0.016	.14	0.019	-0.001, 0.039	.06	0.001	-0.011, 0.013	.93
Std_offset	0.009	0.001, 0.017	.03	0.019	0.002, 0.036	.03	0.003	-0.008, 0.015	.56
Efficiency	-0.025	-0.041, -0.008	.004	-0.043	-0.071, -0.016	.002	-0.012	-0.037, 0.013	.34
Awake_5	0.014	0.006, 0.022	<.001	0.022	0.009, 0.035	.001	0.016	0.005, 0.027	.01
WKD_diff	0.211	0.084, 0.339	.001	0.071	-0.126, 0.268	.48	0.077	-0.144, 0.299	.49
M_insomnia	0.472	0.259, 0.685	<.001	0.381	0.028, 0.734	.04	-0.048	-0.385, 0.289	.78
Dur_10	0.331	0.191, 0.472	<.001	0.340	0.052, 0.627	.02	0.181	-0.051, 0.413	.13

<sup>a</sup>Definitions of sleep features in this table are shown in [Table 1](#).

<sup>b</sup>KCL: King's College London.

<sup>c</sup>CIBER: Centro de Investigación Biomédica en Red.

<sup>d</sup>VUmc: Vrije Universiteit Medisch Centrum.

<sup>e</sup>Slope coefficient estimates for all sleep features.

**Table 6.** Coefficient estimates, 95% confidence intervals, and *P* values from 2-level linear mixed models on the 3 study sites for exploring associations between sleep features<sup>a</sup> and the sleep subscore<sup>b</sup>.

Features	KCL <sup>c</sup>			CIBER <sup>d</sup>			VUmc <sup>e</sup>		
	Coeff. <sup>f</sup>	95% CI	<i>P</i> value	Coeff.	95% CI	<i>P</i> value	Coeff.	95% CI	<i>P</i> value
Av_tst	0.015	-0.021, 0.050	.41	-0.035	-0.116, 0.047	.41	-0.017	-0.070, 0.035	.52
Av_time_bed	0.026	-0.013, 0.066	.19	-0.025	-0.116, 0.065	.58	-0.015	-0.074, 0.043	.61
Deep_pct	-0.027	-0.134, 0.081	.63	-0.196	-0.412, 0.020	.07	-0.191	-0.369, -0.014	.04
Light_pct	-0.024	-0.213, 0.166	.81	0.098	-0.250, 0.445	.58	0.312	0.016, 0.608	.04
REM_pct	-0.116	-0.260, 0.028	.12	-0.037	-0.304, 0.230	.79	-0.169	-0.398, 0.060	.15
NREM_pct	-0.048	-0.194, 0.098	.52	-0.123	-0.389, 0.143	.37	0.125	-0.096, 0.346	.27
Awake_pct	0.165	0.085, 0.245	<.001	0.150	0.020, 0.280	.02	0.049	-0.073, 0.170	.43
Av_onset	0.055	0.008, 0.101	.02	0.075	-0.023, 0.172	.13	0.128	0.054, 0.202	.001
Av_offset	0.102	0.053, 0.150	<.001	0.048	-0.040, 0.135	.29	0.133	0.056, 0.210	.001
REM_L	0.073	-0.255, 0.401	.66	0.146	-0.494, 0.787	.65	-0.171	-0.683, 0.340	.51
Std_tst	0.046	0.022, 0.071	<.001	0.046	-0.002, 0.094	.06	0.043	0.004, 0.082	.03
Std_onset	0.028	-0.015, 0.070	.21	0.089	-0.018, 0.195	.10	0.079	0.020, 0.139	.01
Std_offset	0.046	0.008, 0.084	.02	0.109	0.022, 0.195	.01	0.072	0.016, 0.127	.01
Efficiency	-0.118	-0.196, -0.041	.003	-0.152	-0.280, -0.024	.02	-0.044	-0.162, 0.074	.46
Awake_5	0.047	0.011, 0.083	.01	0.037	-0.022, 0.097	.22	0.013	-0.042, 0.067	.65
WKD_diff	1.169	0.534, 1.804	<.001	0.210	-0.864, 1.284	.70	0.283	-0.830, 1.395	.62
M_insomnia	2.302	1.274, 3.329	<.001	2.777	1.070, 4.485	.001	1.823	0.180, 3.465	.03
Dur_10	1.057	0.387, 1.728	.002	0.576	-0.844, 1.995	.43	0.706	-0.411, 1.823	.22

<sup>a</sup>The definitions of sleep features in this table are shown in [Table 1](#).

<sup>b</sup>The sleep subscore represents the score of subitem 3 in the PHQ-8.

<sup>c</sup>KCL: King's College London.

<sup>d</sup>CIBER: Centro de Investigación Biomédica en Red.

<sup>e</sup>VUmc: Vrije Universiteit Medisch Centrum.

<sup>f</sup>Slope coefficient estimates for all sleep features.

## Discussion

### Principal Findings

In this study, we extracted 18 sleep features through Fitbit data to quantitatively describe participant sleep characteristics in 5 categories (sleep architecture, sleep stability, sleep quality, insomnia, and hypersomnia) associated with the severity of depression. Along with the depressive status worsening, the following changes may be seen in the past 2 weeks: (1) percentage of light/NREM sleep decreased and the percentage of wakefulness during sleep increased (sleep architecture); (2) sleep duration/onset/offset were unstable (sleep stability); (3) reduced sleep efficiency, more awakenings during sleep, and longer weekend catch-up sleep were observed (sleep quality); (4) more days with insomnia were observed (insomnia); (5) more days with hypersomnia were observed (hypersomnia). [Table 4](#) illustrated that our sleep features of these 5 categories could reflect both the participant sleep condition (sleep subscore) and depressive symptom severity (PHQ-8 score) of the past 2 weeks.

### Potential Factors Affecting Associations

We evaluated our models on the research sites separately. From [Table 5](#) and [Table 6](#), we can notice that the associations between sleep features and PHQ-8 score/sleep subscore varied across different sites. Several factors may affect the associations. First, the populations of the 3 sites were significantly different ([Table 2](#)). For example, participants in the CIBER site came from a clinical population and their average age was oldest, so one speculation is that there was less difference between their weekday sleep and weekend sleep for inpatients or people in retirement. Therefore, this may be the reason why the feature of weekend catch-up sleep (*WKD\_diff*) lost significance on the CIBER data. In addition, the reduced significance of features related to sleep onset and offset time on the CIBER site might be related to the regular sleep pattern in CIBER site favors going to bed later, as seen in our previous study [42].

The associations between sleep features and the sleep subscore on the VUmc data ([Table 6](#)) were similar to that in the entire dataset ([Table 4](#)), which demonstrated sleep features have the same ability to capture the sleep condition of participants on the VUmc data. However, the significance of associations

between these sleep features and the PHQ-8 score was reduced in the VUmc data (Table 5). One possible reason is that, as seen on Table 3, the correlation between the sleep subscore and PHQ-8 score in the VUmc data ( $r=.64$ ) was weaker than other 2 study sites (KCL:  $r=.74$  and CIBER:  $r=.78$ ), which may be caused by confounding variables that we did not consider or record in the VUmc population such as medication and occupational status.

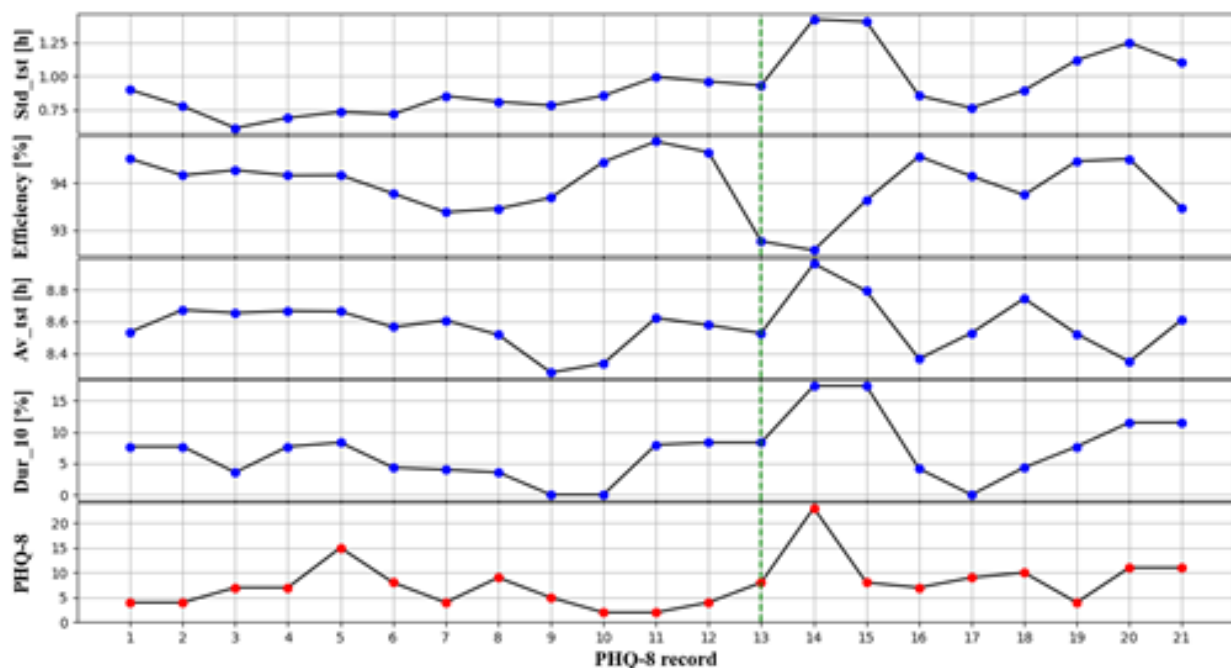
Sample size and heterogeneity of the dataset were other possible factors that may affect results. Table 2 shows that the KCL site had the most PHQ-8 records, whereas VUmc had the least data. As depression manifests itself in distinctive symptoms on different people, it may be difficult to fully explore the associations between sleep and depression on a relatively smaller dataset (VUmc). For example, hypersomnia is specifically related to bipolar patients [7,8]; therefore, if the dataset did not contain enough bipolar patients or bipolar patients were not in depressive episodes when they completed their PHQ-8 records,

it would be hard to find the association between hypersomnia and depression.

### Comparison With Prior Work

Our study has a relatively larger sample size and a longer follow-up duration than previous studies on monitoring depression by using wearable devices and mobile phones [19-21]. Each participant has multiple PHQ-8 records and repeated measurements of sleep, so we can not only explore the relationships between sleep and depression between individuals but also find the associations within individuals by using the linear mixed model. Figure 3 is an example of a possible depression relapse of one participant, showing an obvious increasing trend in PHQ-8 scores at the 13th PHQ-8 record of this participant. We can observe the sleep features in Figure 3 are significantly associated with the PHQ-8 score. This indicates that the sleep features extracted in this paper have the potential to be the biomarkers of depression.

**Figure 3.** The PHQ-8 scores and a select 4 sleep features of one participant with an obvious increasing trend in PHQ-8 score at 13th PHQ-8 record. Descriptions of abbreviations of sleep features in this figure are shown in Table 1.



We also compared our findings with previous studies that used other measurements to assess sleep, such as PSG and sleep questionnaires. Although the sample size, population, measurements, duration of these studies are different, the comparison may help to find more general associations between sleep and depression. Table 7 provides a summary of the comparison. Several longitudinal studies based on sleep questionnaires have shown that insomnia and hypersomnia are both symptoms of depression [6,46], which we found in our research. Kang et al [36] found the weekend catch-up sleep was significantly positively correlated with the severity of depression by analyzing the self-sleep questionnaires of 4553 Korean

adolescents, and this is consistent with the finding in our paper. A sleep report has shown that higher sleep efficiency, more deep sleep, and fewer awakenings after sleep onset represent better sleep quality [32], which is also consistent with the relationships we found between deep sleep percentage, awake percentage, and awakenings (>5 minutes) with sleep subscore. A review showed that according to PSG research, the shortened REM latency and increased percentage of REM sleep are biological markers of depression relapse [9]; however, relationships between depressive symptom severity with REM sleep percentage and REM latency were not significant in our results.

**Table 7.** Summary of the comparisons with previous studies using other measurements to assess sleep.

Type of feature	Findings in previous studies	Consistent <sup>a</sup>	Measurement
Insomnia	Insomnia is significantly related to depression [6].	Yes	Questionnaire
Hypersomnia	Prevalence of hypersomnia is high in depressed patients [46].	Yes	Questionnaire
Weekend catch-up sleep	Weekend catch-up sleep is significantly positively correlated with the severity of depression [36].	Yes	Questionnaire
Deep sleep percentage	More deep sleep represents higher sleep quality [32].	Yes	Questionnaire
Awake percentage, Awakenings (>5 mins)	Fewer awakenings after sleep onset represents better sleep quality [32].	Yes	Questionnaire
Sleep efficiency	Higher sleep efficiency represents better sleep quality [32].	Yes	Questionnaire
REM sleep percentage	Increased REM sleep percentage can be biomarkers of depression [9].	No	Polysomnography
REM <sup>b</sup> latency	Shortened REM latency can be biomarkers of depression [9].	No	Polysomnography

<sup>a</sup>Whether it is consistent with our findings.

<sup>b</sup>REM: rapid eye movement.

## Limitations

Missing data is the major hindrance in our study. For various reasons, there were many missing records of sleep. We set the completion rate of sleep records greater than 85% (12 days) as one of the data inclusion criteria. However, the optimum threshold is unclear, which needs to be further studied in future research. Missingness could also be associated with depressive status and could be a useful marker of relapse of depression; for example, participants may not feel like complying if they are feeling depressed. In future research, we will consider missingness as a potential feature.

Although we adjusted our models for age, gender, education level, and annual income, it is hard to consider all potential confounding variables. For example, some participants with sleep disorders may take sleep medications. Sleep medications have a significant influence on the features of sleep. Unfortunately, there was no daily record of whether the participant took medication. This confounding variable may affect the result.

The data of sleep stages used in this paper were provided by the Fitbit wristband. According to their validation studies, the Fitbit wristband showed promise in detecting sleep-wake states but limitations in other sleep stages estimation [27-29]. This may be the reason the features of REM percentage and REM latency in our paper did not show significant relationships with depressive symptoms. For detecting insomnia, the sleep onset latency (SOL) in the PSG report is a reliable indicator of insomnia, but the Charge 2 and 3 are not able to measure SOL directly. The features related to insomnia in our paper can partially reflect insomnia, but they may be affected by factors (such as work schedules or activities) other than insomnia. Therefore, in future research, we will combine multiple features (such as a late sleep onset time accompanied by a short total sleep time) to determine whether a participant has insomnia and

try to use activity information (eg, steps) provided by Fitbit to approximate SOL. Although there are some limitations of Fitbit data, it provides a means to investigate sleep characteristic in home settings.

In feature extraction, we did not consider the impact of individual circumstances on sleep features. For example, some participants may need to shift work at night, which our features are unable to capture. We will consider the impact of sleep habits and lifestyles on sleep features in the future. Further, we did not explore the impact of individual patterns of depression [47]—for example, the distinction between people with typical and atypical depression who report reduced and increased sleep, respectively, during depressive episodes. In future work, we will explore whether including this dimension improves specificity of our findings.

In this paper, we focused on analyzing the manifestations of depression in sleep characteristics. We will investigate whether these relationships are bidirectional in future research. We only performed bivariate analysis (ie, separately analyzing the association between each feature and the PHQ-8 score). The combination of features and nonlinear relationships was not considered. We will try to apply machine/deep learning models to predict the severity of depression by using sleep features in future research.

## Conclusions

Although consumer wearable devices may not be a substitute for PSG to assess sleep quality accurately, we demonstrated that some derived sleep features extracted from these wearable devices show potential for remote measurement of sleep and consequently can act as a biomarker of depression in real-world settings. These findings may provide the basis for the development of clinical tools that could be used to passively monitor disease state and trajectory with minimal burden on the participant.

## Acknowledgments

The RADAR-CNS project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant 115902. This joint undertaking receives support from the European Union's Horizon 2020 research and innovation program and European Federation of Pharmaceutical Industries and Associations (EFPIA). This communication reflects the views of the RADAR-CNS consortium and neither the Innovative Medicines Initiative nor the European Union and EFPIA are liable for any use that may be made of the information contained herein. Participants in the CIBER site came from the following 4 clinical communities in Spain: Parc Sanitari Sant Joan de Déu Network services, Institut Català de la Salut, Institut Pere Mata, and Hospital Clínico San Carlos. Participant recruitment in Amsterdam was partially accomplished through Hersenonderzoek.nl, a Dutch online registry that facilitates participant recruitment for neuroscience studies [24]. Hersenonderzoek.nl is funded by grant 73305095003 from ZonMw-Memorabel, a project in the context of the Dutch Deltaplan Dementie, Gieskes-Strijbis Foundation, the Alzheimer's Society in the Netherlands, and Brain Foundation Netherlands. This paper represents independent research partially funded by the National Institute for Health Research (NIHR) 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, NIHR, or the Department of Health and Social Care. RB is funded in part by grant MR/R016372/1 from the King's College London Medical Research Council Skills Development Fellowship program funded by the UK Medical Research Council and by grant IS-BRC-1215-20018 from the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London.

## Conflicts of Interest

VAN is an employee of Janssen Research and Development LLC and may own equity in the company.

## References

1. Depression and other common mental disorders: global health estimates. Geneva: World Health Organization; 2017. URL: <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf> [accessed 2021-03-26]
2. 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](https://doi.org/10.1007/s11920-004-0007-y)] [Medline: [15538991](https://pubmed.ncbi.nlm.nih.gov/15538991/)]
3. Lenox-Smith A, Macdonald MTB, Reed C, Tylee A, Peveler R, Quail D, et al. Quality of life in depressed patients in UK primary care: the FINDER study. *Neurol Ther* 2013 Dec;2(1-2):25-42 [FREE Full text] [doi: [10.1007/s40120-013-0006-1](https://doi.org/10.1007/s40120-013-0006-1)] [Medline: [26000214](https://pubmed.ncbi.nlm.nih.gov/26000214/)]
4. 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):S46-S55 [FREE Full text] [doi: [10.1097/01.jom.0000126684.82825.0a](https://doi.org/10.1097/01.jom.0000126684.82825.0a)] [Medline: [15194895](https://pubmed.ncbi.nlm.nih.gov/15194895/)]
5. Mendelson WB, editor. *Human Sleep and Its Disorders*. Berlin: Springer Science & Business Media; 2012.
6. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. *Sleep* 2013 Jul 01;36(7):1059-1068 [FREE Full text] [doi: [10.5665/sleep.2810](https://doi.org/10.5665/sleep.2810)] [Medline: [23814343](https://pubmed.ncbi.nlm.nih.gov/23814343/)]
7. Detre T, Himmelhoch J, Swartzburg M, Anderson CM, Byck R, Kupfer DJ. Hypersomnia and manic-depressive disease. *Am J Psychiatry* 1972 Apr;128(10):1303-1305. [doi: [10.1176/ajp.128.10.1303](https://doi.org/10.1176/ajp.128.10.1303)] [Medline: [4335278](https://pubmed.ncbi.nlm.nih.gov/4335278/)]
8. Thase ME, Himmelhoch JM, Mallinger AG, Jarrett DB, Kupfer DJ. Sleep EEG and DST findings in anergic bipolar depression. *Am J Psychiatry* 1989 Mar;146(3):329-333. [doi: [10.1176/ajp.146.3.329](https://doi.org/10.1176/ajp.146.3.329)] [Medline: [2919689](https://pubmed.ncbi.nlm.nih.gov/2919689/)]
9. Palagini L, Baglioni C, Ciapparelli A, Gemignani A, Riemann D. REM sleep dysregulation in depression: state of the art. *Sleep Med Rev* 2013 Oct;17(5):377-390. [doi: [10.1016/j.smrv.2012.11.001](https://doi.org/10.1016/j.smrv.2012.11.001)] [Medline: [23391633](https://pubmed.ncbi.nlm.nih.gov/23391633/)]
10. Riemann D, Berger M, Voderholzer U. Sleep and depression—results from psychobiological studies: an overview. *Biol Psychol* 2001;57(1-3):67-103. [doi: [10.1016/s0301-0511\(01\)00090-4](https://doi.org/10.1016/s0301-0511(01)00090-4)] [Medline: [11454435](https://pubmed.ncbi.nlm.nih.gov/11454435/)]
11. Iber C, Ancoli-Israel S, Chesson A, Quan S. *AASM Manual for the Scoring of Sleep and Associated Events*. Darien: American Academy of Sleep Medicine; 2007.
12. Sánchez-Ortuño MM, Edinger JD, Means MK, Almirall D. Home is where sleep is: an ecological approach to test the validity of actigraphy for the assessment of insomnia. *J Clin Sleep Med* 2010 Feb 15;6(1):21-29 [FREE Full text] [Medline: [20191934](https://pubmed.ncbi.nlm.nih.gov/20191934/)]
13. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989 May;28(2):193-213. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](https://pubmed.ncbi.nlm.nih.gov/2748771/)]
14. Moore CM, Schmiede SJ, Matthews EE. Actigraphy and sleep diary measurements in breast cancer survivors: discrepancy in selected sleep parameters. *Behav Sleep Med* 2015;13(6):472-490 [FREE Full text] [doi: [10.1080/15402002.2014.940108](https://doi.org/10.1080/15402002.2014.940108)] [Medline: [25117292](https://pubmed.ncbi.nlm.nih.gov/25117292/)]
15. Beattie Z, Oyang Y, Statan A, Ghoreyshi A, Pantelopoulos A, Russell A, et al. Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals. *Physiol Meas* 2017 Oct 31;38(11):1968-1979. [doi: [10.1088/1361-6579/aa9047](https://doi.org/10.1088/1361-6579/aa9047)] [Medline: [29087960](https://pubmed.ncbi.nlm.nih.gov/29087960/)]

16. Van de Water ATM, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography—a systematic review. *J Sleep Res* 2011 Mar;20(1 Pt 2):183-200 [FREE Full text] [doi: [10.1111/j.1365-2869.2009.00814.x](https://doi.org/10.1111/j.1365-2869.2009.00814.x)] [Medline: [20374444](https://pubmed.ncbi.nlm.nih.gov/20374444/)]
17. Zhang Y, Yang Z, Lan K. Sleep stage classification using bidirectional lstm in wearable multi-sensor systems. 2019 Presented at: IEEE INFOCOM—IEEE Conference on Computer Communications Workshops; 2019; Paris p. 443-448. [doi: [10.1109/infcomw.2019.8845115](https://doi.org/10.1109/infcomw.2019.8845115)]
18. Zhang Y, Yang Z, Zhang Z. Breathing disorder detection using wearable electrocardiogram and oxygen saturation. *Proc 16th ACM Conf Emb Netw Sensor Syst* 2018:313-314. [doi: [10.1145/3274783.3275159](https://doi.org/10.1145/3274783.3275159)]
19. Miwa H, Sasahara S, Matsui T. Roll-over detection and sleep quality measurement using a wearable sensor. *Annu Int Conf IEEE Eng Med Biol Soc* 2007:1507-1510. [doi: [10.1109/IEMBS.2007.4352587](https://doi.org/10.1109/IEMBS.2007.4352587)] [Medline: [18002253](https://pubmed.ncbi.nlm.nih.gov/18002253/)]
20. Mark G, Czerwinski M, Iqbal S. Workplace indicators of mood: behavioral and cognitive correlates of mood among information workers. *Proc 6th Int Conf on Dig Health* 2016:29-36. [doi: [10.1145/2896338.2896360](https://doi.org/10.1145/2896338.2896360)]
21. Demasi O, Aguilera A, Recht B. Detecting change in depressive symptoms from daily wellbeing questions, personality, and activity. *IEEE* 2016:1. [doi: [10.1109/wh.2016.7764552](https://doi.org/10.1109/wh.2016.7764552)]
22. Khouli S, Garzón-Rey J, Aguilo J. Remote Assessment of Disease and Relapse—Central Nervous System (RADAR-CNS). *Transact Mach Learn Artif Intell* 2017 Aug 31;5(4):1. [doi: [10.14738/tmlai.54.3293](https://doi.org/10.14738/tmlai.54.3293)]
23. 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 Feb 18;19(1):72 [FREE Full text] [doi: [10.1186/s12888-019-2049-z](https://doi.org/10.1186/s12888-019-2049-z)] [Medline: [30777041](https://pubmed.ncbi.nlm.nih.gov/30777041/)]
24. Hersenonderzoek.nl. URL: <https://herenonderzoek.nl> [accessed 2020-12-07]
25. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeek D, 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](https://doi.org/10.2196/11734)] [Medline: [31373275](https://pubmed.ncbi.nlm.nih.gov/31373275/)]
26. Kroenke K, Strine TW, Spitzer RL, Williams JBW, 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](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
27. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2 compared with polysomnography in adults. *Chronobiol Int* 2018 Apr;35(4):465-476. [doi: [10.1080/07420528.2017.1413578](https://doi.org/10.1080/07420528.2017.1413578)] [Medline: [29235907](https://pubmed.ncbi.nlm.nih.gov/29235907/)]
28. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband fitbit models in assessing sleep: systematic review and meta-analysis. *J Med Internet Res* 2019 Nov 28;21(11):e16273 [FREE Full text] [doi: [10.2196/16273](https://doi.org/10.2196/16273)] [Medline: [31778122](https://pubmed.ncbi.nlm.nih.gov/31778122/)]
29. Liang Z, Chapa-Martell MA. Accuracy of fitbit wristbands in measuring sleep stage transitions and the effect of user-specific factors. *JMIR Mhealth Uhealth* 2019 Jun 06;7(6):e13384 [FREE Full text] [doi: [10.2196/13384](https://doi.org/10.2196/13384)] [Medline: [31172956](https://pubmed.ncbi.nlm.nih.gov/31172956/)]
30. Aluoja A, Leinsalu M, Shlik J, Vasar V, Luuk K. Symptoms of depression in the Estonian population: prevalence, sociodemographic correlates and social adjustment. *J Affect Disord* 2004 Jan;78(1):27-35. [doi: [10.1016/s0165-0327\(02\)00179-9](https://doi.org/10.1016/s0165-0327(02)00179-9)] [Medline: [14672794](https://pubmed.ncbi.nlm.nih.gov/14672794/)]
31. Akhtar-Danesh N, Landeen J. Relation between depression and sociodemographic factors. *Int J Ment Health Syst* 2007 Sep 04;1(1):4 [FREE Full text] [doi: [10.1186/1752-4458-1-4](https://doi.org/10.1186/1752-4458-1-4)] [Medline: [18271976](https://pubmed.ncbi.nlm.nih.gov/18271976/)]
32. Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, et al. National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health* 2017 Feb;3(1):6-19. [doi: [10.1016/j.sleh.2016.11.006](https://doi.org/10.1016/j.sleh.2016.11.006)] [Medline: [28346153](https://pubmed.ncbi.nlm.nih.gov/28346153/)]
33. Walker MP. Sleep-dependent memory processing. *Harv Rev Psychiatry* 2008;16(5):287-298. [doi: [10.1080/10673220802432517](https://doi.org/10.1080/10673220802432517)] [Medline: [18803104](https://pubmed.ncbi.nlm.nih.gov/18803104/)]
34. Rasch B, Born J. About sleep's role in memory. *Physiol Rev* 2013 Apr;93(2):681-766 [FREE Full text] [doi: [10.1152/physrev.00032.2012](https://doi.org/10.1152/physrev.00032.2012)] [Medline: [23589831](https://pubmed.ncbi.nlm.nih.gov/23589831/)]
35. Manni R. Rapid eye movement sleep, non-rapid eye movement sleep, dreams, and hallucinations. *Curr Psychiatry Rep* 2005 Jun;7(3):196-200. [doi: [10.1007/s11920-005-0053-0](https://doi.org/10.1007/s11920-005-0053-0)] [Medline: [15935133](https://pubmed.ncbi.nlm.nih.gov/15935133/)]
36. Kang S, Lee YJ, Kim SJ, Lim W, Lee H, Park Y, et al. Weekend catch-up sleep is independently associated with suicide attempts and self-injury in Korean adolescents. *Compr Psychiatry* 2014 Feb;55(2):319-325. [doi: [10.1016/j.comppsy.2013.08.023](https://doi.org/10.1016/j.comppsy.2013.08.023)] [Medline: [24267542](https://pubmed.ncbi.nlm.nih.gov/24267542/)]
37. Liu X, Zhao Z, Jia C, Buysse DJ. Sleep patterns and problems among chinese adolescents. *Pediatrics* 2008 Jun;121(6):1165-1173. [doi: [10.1542/peds.2007-1464](https://doi.org/10.1542/peds.2007-1464)] [Medline: [18519486](https://pubmed.ncbi.nlm.nih.gov/18519486/)]
38. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington: American Psychiatric Publishers; 2013.
39. Tam EM, Lam RW, Robertson HA, Stewart JN, Yatham LN, Zis AP. Atypical depressive symptoms in seasonal and non-seasonal mood disorders. *J Affect Disord* 1997 Jun;44(1):39-44. [doi: [10.1016/s0165-0327\(97\)01447-x](https://doi.org/10.1016/s0165-0327(97)01447-x)] [Medline: [9186801](https://pubmed.ncbi.nlm.nih.gov/9186801/)]
40. Farhan AA, Yue C, Morillo R. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. 2016 Presented at: 2016 IEEE Wireless Health (WH); 2016; Bethesda p. 1-8. [doi: [10.1109/wh.2016.7764553](https://doi.org/10.1109/wh.2016.7764553)]

41. Singer JD, Willett JB, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford: Oxford University Press; 2003.
42. Sun S, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, RADAR-CNS Consortium. Using smartphones and wearable devices to monitor behavioral changes during COVID-19. *J Med Internet Res* 2020 Sep 25;22(9):e19992 [FREE Full text] [doi: [10.2196/19992](https://doi.org/10.2196/19992)] [Medline: [32877352](https://pubmed.ncbi.nlm.nih.gov/32877352/)]
43. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982 Dec;38(4):963-974. [Medline: [7168798](https://pubmed.ncbi.nlm.nih.gov/7168798/)]
44. Box GEP, Cox DR. An analysis of transformations. *J Royal Stat Soc B* 2018 Dec 05;26(2):211-243. [doi: [10.1111/j.2517-6161.1964.tb00553.x](https://doi.org/10.1111/j.2517-6161.1964.tb00553.x)]
45. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc B* 2018 Dec 05;57(1):289-300. [doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)]
46. Kaplan KA, Harvey AG. Hypersomnia across mood disorders: a review and synthesis. *Sleep Med Rev* 2009 Aug;13(4):275-285. [doi: [10.1016/j.smrv.2008.09.001](https://doi.org/10.1016/j.smrv.2008.09.001)] [Medline: [19269201](https://pubmed.ncbi.nlm.nih.gov/19269201/)]
47. Brailean A, Curtis J, Davis K, Dregan A, Hotopf M. Characteristics, comorbidities, and correlates of atypical depression: evidence from the UK Biobank Mental Health Survey. *Psychol Med* 2019 May 02;50(7):1129-1138. [doi: [10.1017/s0033291719001004](https://doi.org/10.1017/s0033291719001004)]

## Abbreviations

**CIBER:** Centro de Investigación Biomédica en Red

**EFPIA:** European Federation of Pharmaceutical Industries and Associations

**KCL:** King's College London

**NHS:** National Health Service

**NIHR:** National Institute for Health Research

**NREM:** non-REM

**PHQ-8:** Patient Health Questionnaire 8-item

**PSG:** polysomnography

**PSQI:** Pittsburgh Sleep Quality Index

**RADAR-CNS:** Remote Assessment of Disease and Relapse–Central Nervous System

**RADAR-MDD:** Remote Assessment of Disease and Relapse–Major Depressive Disorder

**REM:** rapid eye movement

**SOL:** sleep onset latency

**VUmc:** Vrije Universiteit Medisch Centrum

**WHO:** World Health Organization

*Edited by L Buis; submitted 26.09.20; peer-reviewed by N Jacobson, Z Liang; comments to author 25.11.20; revised version received 07.12.20; accepted 03.02.21; published 12.04.21.*

*Please cite as:*

Zhang Y, Folarin AA, Sun S, Cummins N, Bendayan R, Ranjan Y, Rashid Z, Conde P, Stewart C, Laiou P, Matcham F, White KM, Lamers F, Siddi S, Simblett S, Myin-Germeyns I, Rintala A, Wykes T, Haro JM, Penninx BWJH, Narayan VA, Hotopf M, Dobson RJB, RADAR-CNS Consortium

*Relationship Between Major Depression Symptom Severity and Sleep Collected Using a Wristband Wearable Device: Multicenter Longitudinal Observational Study*

*JMIR Mhealth Uhealth* 2021;9(4):e24604

URL: <https://mhealth.jmir.org/2021/4/e24604>

doi: [10.2196/24604](https://doi.org/10.2196/24604)

PMID: [33843591](https://pubmed.ncbi.nlm.nih.gov/33843591/)

©Yuezhou Zhang, Amos A Folarin, Shaoxiong Sun, Nicholas Cummins, Rebecca Bendayan, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Petroula Laiou, Faith Matcham, Katie M White, Femke Lamers, Sara Siddi, Sara Simblett, Inez Myin-Germeyns, Aki Rintala, Til Wykes, Josep Maria Haro, Brenda WJH Penninx, Vaibhav A Narayan, Matthew Hotopf, Richard JB Dobson, RADAR-CNS Consortium. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 12.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Associations of Physical Activity Level and Variability With 6-Month Weight Change Among 26,935 Users of Connected Devices: Observational Real-Life Study

Douae El Fatouhi<sup>1</sup>, MSc; Lidia Delrieu<sup>2</sup>, PhD; Catherine Goetzinger<sup>3,4</sup>, MSc; Laurent Malisoux<sup>3</sup>, PhD; Aurélie Affret<sup>1</sup>, PhD; David Campo<sup>5</sup>, PhD; Guy Fagherazzi<sup>1,3</sup>, PhD

<sup>1</sup>Center of Research in Epidemiology and Population Health, UMR 1018 INSERM, Institut Gustave Roussy, Paris-Sud Paris-Saclay University, Villejuif, France

<sup>2</sup>Residual Tumor & Response to Treatment Laboratory (RT2Lab), U932 Immunity and Cancer, INSERM, Institut Curie, Paris, France

<sup>3</sup>Department of Population Health, Luxembourg Institute of Health, Strassen, Luxembourg

<sup>4</sup>Faculty of Science, Technology and Medicine, University of Luxembourg, Luxembourg, Luxembourg

<sup>5</sup>Withings, Issy-les-Moulineaux, France

**Corresponding Author:**

Guy Fagherazzi, PhD

Department of Population Health

Luxembourg Institute of Health

1 A-B Rue Thomas Edison

Strassen, L-1445

Luxembourg

Phone: 352 26 97 04 57

Email: [guy.fagherazzi@lih.lu](mailto:guy.fagherazzi@lih.lu)

## Abstract

**Background:** Physical activity (PA) is a modifiable lifestyle factor that can be targeted to increase energy expenditure and promote weight loss. However, the amount of PA required for weight loss remains inconsistent. Wearable activity trackers constitute a valuable opportunity to obtain objective measurements of PA and study large populations in real-life settings.

**Objective:** We aim to study the associations of initial device-assessed PA characteristics (average step counts and step count variability) and their evolution with 6-month weight change.

**Methods:** We analyzed data from 26,935 Withings-connected device users (wearable activity trackers and digital scales). To assess the initial PA characteristics and their 6-month changes, we used data recorded during the first and sixth 30-day periods of activity tracker use. For each of these periods, we used the monthly mean of daily step values as a proxy for PA level and derived the monthly coefficient of variation (CV) of daily step values to estimate PA level variability. Associations between initial PA characteristics and 6-month weight change were assessed using multivariable linear regression analyses controlled for age, sex, blood pressure, heart rate, and the predominant season. Restricted cubic spline regression was performed to better characterize the continuous shape of the associations between PA characteristics and weight change. Secondary analyses were performed by analyzing the 6-month evolution of PA characteristics in relation to weight change.

**Results:** Our results revealed that both a greater PA level and lower PA level variability were associated with weight loss. Compared with individuals who were initially in the sedentary category (<5000 steps/day), individuals who were low active (5000-7499 steps/day), somewhat active (7500-9999 steps/day), and active ( $\geq 10,000$  steps/day) had a 0.21-kg, a 0.52-kg, and a 1.17-kg greater decrease in weight, respectively (95% CI -0.36 to -0.06, -0.70 to -0.33, and -1.42 to -0.93, respectively). Compared with users whose PA level CV was  $>63\%$ , users whose PA level CV ranged from 51% to 63%, 40% to 51%, and was  $\leq 40\%$ , had a 0.19-kg, a 0.23-kg, and a 0.33-kg greater decrease in weight, respectively (95% CI -0.38 to -0.01, -0.41 to -0.04, and -0.53 to -0.13, respectively). We also observed that each 1000 steps/day increase in PA level over the 6-month follow-up was associated with a 0.26-kg (95% CI -0.29 to -0.23) decrease in weight. No association was found between the 6-month changes in PA level variability and weight change.

**Conclusions:** Our results add to the current body of knowledge that health benefits can be observed below the 10,000 steps/day threshold and suggest that not only increased mean PA level but also greater regularity of the PA level may play important roles in short-term weight loss.



**KEYWORDS**

connected devices; Withings; physical activity; step count; wearable activity trackers; digital health; free-living; weight loss; digital scale; mobile phone

## Introduction

### Background

The prevalence of overweight and obesity is continuously increasing; in 2016, 52% of adults worldwide were overweight or obese [1]. Excess body weight, including overweight and obesity, has been associated with an increased risk of several chronic diseases such as cardiovascular diseases, type 2 diabetes, and some types of cancer [2,3] and thus represents a major public health priority. Therefore, weight loss represents a key leverage to prevent or manage numerous obesity-related conditions. Individuals who are overweight and obese are strongly encouraged to attempt at least clinically significant weight loss ( $\geq 5\%$  of baseline body weight) [4,5].

Studies have shown that physical activity (PA) has multiple health-related benefits, especially among those who are overweight and obese [6-11]. PA recommendations promote the accumulation of at least 150 minutes per week of moderate-to-vigorous PA for maintaining and improving health. These PA guidelines can also be expressed in terms of steps per day. The most popular translation of existing PA guidelines into steps per day equivalents stated that adults are encouraged to accumulate 10,000 steps per day, 3 to 4 days a week [12,13]. However, the scientific literature on the minimal requirement of PA for weight loss remains to be inconsistent. Although some previous studies suggest that engaging in a PA program following the current public health guidelines may result in no weight loss or modest weight loss in individuals who are overweight and obese [14,15], others have shown that substantial weight loss could still occur if the overall amount of PA far exceeds the minimum recommended level of 150 minutes per week in moderate-to-vigorous PA intensity, particularly by surpassing at least a PA volume of 225 minutes per week [13-16]. In addition, most studies have evaluated the effects of the total amount of PA, but little is known about the best combination of total volume and regularity of PA practice over time.

Monitoring PA over a long period is challenging, especially in large and real-life observational cohort studies. Studies on the associations between PA and weight loss are mostly based on self-reported assessments of both variables. It is well known that self-reported data are prone to social desirability and memory recall bias [17,18]. This often results in overreporting of typical PA habits [19]. Consumer-based wearable activity trackers therefore constitute a valuable opportunity to obtain, over large periods, more objective and precise measurements of PA characteristics. Wearable activity trackers are small, affordable devices that are largely commercially available and allow individuals to objectively and continuously monitor their PA levels by providing real-time feedback [20]. In addition,

self-tracking of weight using connected Wi-Fi scales represents a better alternative than self-reported weight values, as it avoids some bias. Data on weight self-monitoring collected from smart scales can also open new research perspectives, including the study of weight loss and weight management among larger populations of real-world users [21,22]. However, these data remain to be underexploited for clinical and epidemiological research purposes.

### Objectives

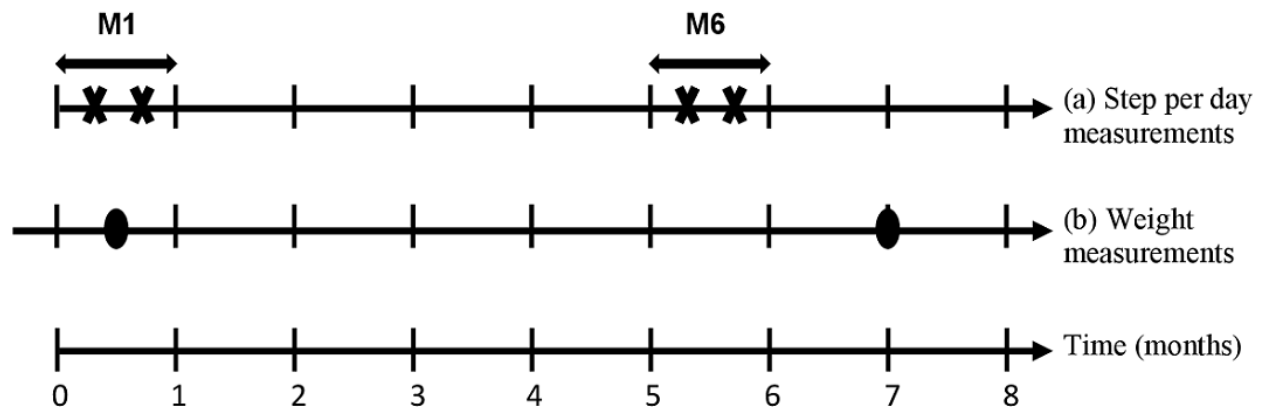
Therefore, based on a large international population of more than 25,000 users of connected activity trackers and digital scales, we first aim to study the associations of device-assessed PA characteristics (mean PA level and variability in PA level) with weight change over a 6-month period. The secondary aim of this study is to analyze the associations between the evolution of PA characteristics over time and the 6-month weight change.

## Methods

### Study Sample and Available Data

The population was initially composed of a sample of 35,841 highly connected Withings customers, those who had purchased at least three Withings-connected devices: a pulse activity tracker [23,24], a body weighing scale [25], and a BP-800 blood pressure monitor [26] (Withings SA). Customer data, including sex, age, number of steps per day, body weight, heart rate, and blood pressure, were provided by Withings and were collected between October 2009 and April 2016. Data were preprocessed before analysis, as we excluded some outliers from daily data on weight and steps with unlikely values as follows: (1) a lower threshold of 500 steps per day was used to denote valid days with daily step count data [27,28], but the application of this criterion did not result in any deletion of data as the minimum daily step value was 501 steps per day; (2) daily step values exceeding the 99<sup>th</sup> percentile of the steps per day distribution were removed (99<sup>th</sup> percentile=22,414 steps/day; 151,937 measurements removed); and (3) weight values lower than 45 kg and above 200 kg were excluded (12,403 measurements removed). In addition, individuals were excluded if they did not (1) have at least two step measurements within the first and sixth 30-day period of activity tracker use ( $n=4733$ ) and (2) have at least one weight measurement within the first 30-day period of activity tracker use and 6 months after ( $n=4173$ ). **Figure 1** shows the inclusion criteria of this study. The final study sample comprised 26,935 individuals (23,580 males and 3355 females). Consent for participation in this study was obtained electronically when users created their Withings account and accepted the treatment of personal data anonymously by Withings for research purposes.

**Figure 1.** Follow-up and inclusion criteria of the study (N=26,935). (a) Step per day measurements were obtained through the Withings Pulse activity tracker. (b) Weight was assessed by users on their Withings weight scale. M1: first 30-day period of activity tracker use; M6: sixth 30-day period of activity tracker use.



#### Inclusion:

- ✕ Step: having  $\geq 2$  measures within M1 and M6 (average number of measures 25.0, SD 6.8 and 24.2, SD 7.4)
- Weight: having  $\geq 1$  measure within M1 and 6 months after (average number of measures 16.4, SD 9.0 and 14.7, SD 9.7)

## PA Assessment

### Monitoring of PA

The Withings Pulse activity tracker is a commercially available device that can be worn on the wrist or clipped onto a belt or clothing. This activity tracker monitors physical behaviors including the total number of steps taken per day. In addition, it provides the user with real-time feedback on their health metrics through the tracker itself or via an app on the user's smartphone or tablet with internet access. The wearable device and the app are synchronized, which allows the data recorded by the activity tracker to be instantly transferred from the device to the app via Bluetooth, thus enabling long-term data tracking by users themselves or, for instance, by researchers or health professionals. Wearable activity trackers are considered to be an accurate and objective assessment method that permits PA monitoring [29]. In particular, the Withings Pulse activity tracker has been shown to be reliable and valid for measuring the number of steps taken by healthy subjects under free-living conditions [24,29-31].

### Assessment of PA Characteristics

To assess the initial PA characteristics and the 6-month change in PA characteristics, we used data recorded during the first and sixth 30-day periods of activity tracker use (named *M1* and *M6* in the following text, respectively). We computed for each of these periods: (1) the mean daily steps and (2) the coefficient of variation (CV) of daily steps. The information provided by the activity tracker that we used in this study was the total daily steps, which is one PA parameter that reflects the user's PA behavior, and especially bipedal locomotion. This metric is considered a reliable proxy of the overall volume of the PA performed by a subject [13]. Averaging it across a 30-day period gives us an indication of the individuals' PA levels. The CV of

daily steps is the ratio of the SD of daily steps, estimated using data recorded during a 30-day period, to the corresponding mean multiplied by 100. The CV is an indicator of the variability in PA levels and is considered here as a proxy of the regularity of the level of PA practice. A low CV value indicates low variability and therefore a high regularity of the PA level within the 30-day period used to compute the CV. Initial PA characteristics included mean PA level, which corresponds to mean daily steps during *M1*, and regularity or variability in PA level, which will be referred to as PA level CV. The 6-month change in PA characteristics included change in mean PA level and change in the variability of PA level.

### Assessment of Weight and Weight Change

We focused on the weight change of users over a 6-month period of use of a consumer-grade activity tracker in a real-life setting. Body weight, expressed in kg, was assessed using a connected Withings digital scale. The analyses only included individuals with at least one weight measurement within the *M1* (Figure 1). The weight value closest to the beginning of *M1* was used as initial body weight. Users also had to have at least one weight measurement after seven 30-day periods of activity tracker use. The weight value closest to the end of the seventh 30-day period of activity tracker use was used as the final body weight. The 6-month weight change was calculated as final weight–initial weight. This anthropometric parameter constituted the main outcome of interest.

### Statistical Analysis

In the primary analysis, multivariable linear regression models were used to explore the relationship between initial PA characteristics and 6-month weight change. PA characteristics included the mean PA level and variability in PA levels during *M1*. We first analyzed both the continuous mean PA level (per

1000 steps/day) and PA level CV (per 10%). We then studied a 4-category PA level classification based on Tudor-Locke classification [12,13] and the PA level CV divided into 4 quartile groups. The 4-category PA level classification consisted of sedentary (<5000 steps/day), low active (5000-7499 steps/day), somewhat active (7500-9999 steps/day), and active ( $\geq 10,000$  steps/day). The sedentary category was used as the reference in the models for PA level, and for the variability in PA level, the highest quartile of CV was used as the reference category (high variability, ie, low regularity in PA level).

Multivariable models were adjusted for age category (18-30 years, 31-40 years, 41-50 years, 51-60 years, and >60 years), sex, diastolic blood pressure, systolic blood pressure, heart rate, and the predominant season during the 6-month follow-up. Blood pressure and heart rate values were assessed by computing the average of all the measures available within M1 and 6 months before M1. Study participants were categorized according to the tertile distribution of blood pressure and heart rate. Missing values for these variables were greater than 5% of the study sample; thus, missing categories were created. Indicators of PA level and PA level CV were simultaneously included in the multivariable models.

Tests for linear trends were conducted by assigning the median value of mean PA level and PA level CV to each category or quartile group and modeling this value as a continuous variable.

Multivariable restricted cubic spline regression was used to better characterize the continuous shape of the associations between mean PA level, regularity of the level, and 6-month weight change. The reference values for estimating the difference in weight change and 95% CIs were chosen as the minimum and maximum values for mean PA level and PA level CV, and 3 knots were used (the 25th, 50th, and 75th percentiles of the mean PA level and the PA level CV distributions). Therefore, restricted cubic spline regression provides us with a graphical presentation in which the y-axis represents the difference in the 6-month weight change associated with any value of mean PA level (or PA level CV) when compared with the minimum value of mean PA level (or the maximum value of PA level CV).

In a secondary analysis, both univariable and multivariable linear regression models were used to analyze the associations between the 6-month change in PA characteristics and the 6-month weight change. Changes in PA characteristics included changes in mean PA levels and in PA level variability. The change in mean PA level was studied using the continuous change (per 1000 steps/day), the change in the 4-category step-defined PA classification, and a 5-category change in mean PA level. This 5-category classification consisted of a decrease in mean PA level of more than 3000 steps per day, decrease between 3000 and 1000 steps per day, change in mean PA level between -1000 and +1000 steps per day, increase in mean PA level between 1000 and 3000 steps per day, and an increase of more than 3000 steps per day. Change in the variability of PA level included the continuous change (ie, the difference in PA

level CV between M6 and M1, per 10%) and the change in quartile-based categories of PA level CV. We first defined quartile-based cut-offs at M1 and used them to categorize PA level CV at M1 and M6, and the 6-month changes between these categories were then assessed.

Statistical analyses were performed using SAS software (version 9.4; SAS Institute), and two-sided *P* values <.05 were considered statistically significant. Graphs were generated using SAS and Python (version 3.5) software.

## Results

### Characteristics of the Study Population

The characteristics of the study population are displayed in Table 1 as means and standard deviations for continuous variables and as numbers and percentages for categorical variables. Individuals were described according to the initial PA characteristics based on the mean PA level and variability in PA level (PA level CV). Most users included in our sample were male (23,580/26,935, 87.54%) and aged between 41 years and 60 years (16,439/26,935, 61.03%). The initial weight in our study population had a mean value of 88.9 kg (SD 18.9). Means of PA level and PA level CV assessed during M1 were 5940.4 (SD 2929.8) steps per day and 53.4% (SD 18.9%), respectively, based on a median of 28.0 (range 2-30; Q1-Q3 23-30, IQR 7.0) step measurements per user. During the sixth 30-day period of activity tracker use (M6), the mean PA level and PA level CV were 6084.7 (SD 2977.7) steps per day and 52.1% (SD 18.6%), respectively, based on a median of 28.0 (range 2-30; Q1-Q3 21-30, IQR 9.0) step measurements per user. More than 75% of the study population had at least 20 step measurements during M1 and M6 (first quartile Q1=23 step measurements during M1 and Q1=21 step measurements during M6). Individuals in the higher step-defined PA categories tended to decrease their PA level and increase their PA level CV during the follow-up, whereas those in the lower categories tended to increase their PA level and decrease their PA level CV during follow-up. We observed similar trends in the PA level CV quartile groups, that is, individuals in the lower PA level CV quartile groups (more regular) tended to decrease their PA level and increase their PA level CV during the follow-up, whereas those in the higher quartile groups (less regular) tended to increase their PA level and decrease their PA level CV during the follow-up. Mean systolic blood pressure, diastolic blood pressure, and heart rate was 127.4 (SD 11.4) mm Hg, 79.2 (SD 8.2) mm Hg, and 70.6 (SD 10.0) bpm, respectively. Health parameters tended to be better in the higher step-defined PA categories than in the lowest category, including lower body weight, blood pressure, heart rate, and higher weight loss. However, this seemed to be the case only for systolic blood pressure and weight loss when comparing individuals in the lower quartile groups of PA level CV (more regular) with those in the highest quartile group (less regular). Figure 2 shows real-world examples of PA monitoring within the first 30-day period of the use of a wearable activity tracker for 4 users with different PA characteristics.

**Table 1.** Characteristics of the study population (N=26,935).

Variable	Overall (N=26,935)	4-category step-defined physical activity level (steps/day, M1 <sup>a</sup> )				Physical activity level variability (CV <sup>b</sup> , M1)			
		Sedentary (<5000 steps/day; n=11,232)	Low active (5000- 7500; n=8266)	Somewhat active (7500- 10,000; n=4837)	Active (≥10,000; n=2600)	Quartile 1 (≤40%; high regularity; n=6733)	Quartile 2 (40%- 51%; n=6734)	Quartile 3 (51%- 63%; n=6735)	Quartile 4 (>63%; low regularity; n=6733)
<b>Age category (years), n (%)</b>									
18-30	553 (2.1)	206 (1.8)	184 (2.2)	103 (2.1)	60 (2.3)	139 (2.1)	150 (2.2)	145 (2.2)	119 (1.8)
31-40	3556 (13.2)	1298 (11.6)	1221 (14.8)	692 (14.3)	345 (13.3)	817 (12.1)	977 (14.5)	923 (13.7)	839 (12.5)
41-50	8214 (30.5)	3095 (27.6)	2678 (32.4)	1621 (33.5)	820 (31.5)	2058 (30.6)	2050 (30.5)	2112 (31.4)	1994 (29.5)
51-60	8225 (30.5)	3427 (30.5)	2471 (29.9)	1503 (31.1)	824 (31.7)	2007 (29.8)	2068 (30.7)	2032 (30.2)	2118 (31.5)
>60	6387 (23.7)	3206 (28.5)	1712 (20.7)	918 (19.0)	551 (21.2)	1712 (25.4)	1489 (22.1)	1523 (22.5)	1663 (24.7)
<b>Sex, n (%)</b>									
Male	23,580 (87.5)	9589 (85.4)	7406 (89.6)	4281 (88.5)	2304 (88.6)	5854 (86.9)	5947 (88.3)	5965 (88.6)	5814 (86.4)
Female	3355 (12.5)	1643 (14.6)	860 (10.4)	556 (11.5)	296 (11.4)	879 (13.1)	787 (11.7)	770 (11.4)	919 (13.6)
Baseline weight (kg), mean (SD)	88.9 (18.9)	91.7 (20.7)	88.5 (17.7)	85.9 (16.7)	84.1 (15.8)	87.5 (18.6)	89.6 (19.1)	89.7 (18.8)	88.9 (18.8)
6-month weight change (kg) <sup>c</sup> , mean (SD)	-1.6 (5.4)	-1.3 (5.5)	-1.6 (5.1)	-1.9 (5.3)	-2.6 (5.5)	-2.0 (5.6)	-1.7 (5.4)	-1.6 (5.4)	1.2 (5.0)
Physical activi- ty level during M1 (steps/day), mean (SD)	5940.4 (2929.8)	3276.6 (1115.5)	6161.6 (713.2)	8593.0 (714.8)	11809.8 (1621.5)	7666.9 (3475.4)	6266.3 (2694.0)	5421.2 (2394.3)	4407.2 (1887.5)
Physical activi- ty level CV dur- ing M1 (%), mean (SD)	53.4 (18.9)	60.7 (20.7)	53.3 (16.0)	45.7 (12.6)	36.5 (11.3)	32.7 (6.3)	45.6 (3.0)	56.4 (3.6)	78.9 (14.8)
Number of step measurements during M1 <sup>d</sup> , median (IQR)	28.0 (23.0- 30.0)	25.0 (18.0- 29.0)	29.0 (25.0- 30.0)	29.0 (27.0- 30.0)	29.0 (26.0- 30.0)	29.0 (26.0- 30.0)	29.0 (25.0- 30.0)	28.0 (23.0- 30.0)	25.0 (19.0- 29.0)
Physical activi- ty level during M6 <sup>e</sup> (steps/day), mean (SD)	6084.7 (2977.7)	4203.5 (2068.8)	6232.7 (2139.8)	7984.8 (2394.6)	10206.4 (3013.2)	7364.3 (3314.5)	6341.2 (2852.6)	5705.5 (2641.9)	4928.1 (2486.5)
Physical activi- ty level CV dur- ing M6 (%), mean (SD)	52.1 (18.6)	57.3 (20.0)	51.2 (16.6)	46.9 (15.5)	41.9 (16.0)	42.5 (15.2)	49.0 (15.3)	54.4 (17.0)	62.4 (20.5)
Number of step measurements during M6 <sup>d</sup> , median (IQR)	28.0 (21.0- 30.0)	26.0 (18.0- 29.0)	28.0 (22.0- 30.0)	29.0 (24.0- 30.0)	28.0 (24.5- 30.0)	29.0 (23.0- 30.0)	28.0 (22.0- 30.0)	27.0 (21.0- 30.0)	26.0 (19.0- 29.0)
Change in phys- ical activity lev- el (steps/day) <sup>c</sup> , mean (SD)	144.4 (2305.5)	926.9 (1948.0)	71.2 (2089.0)	-608.2 (2366.6)	-1603.4 (2759.4)	-302.7 (2372.9)	74.9 (2287.5)	284.3 (2194.7)	520.9 (2284.2)
Change in phys- ical activity lev- el CV (%) <sup>c</sup> , mean (SD)	-1.3 (20.4)	-3.4 (24.2)	-2.1 (17.9)	1.2 (15.6)	5.4 (15.2)	9.8 (15.5)	3.4 (15.4)	-2.0 (17.0)	-16.5 (22.8)

Variable	Overall (N=26,935)	4-category step-defined physical activity level (steps/day, M1 <sup>a</sup> )				Physical activity level variability (CV <sup>b</sup> , M1)			
		Sedentary ( $<5000$ steps/day; n=11,232)	Low active (5000- 7500; n=8266)	Somewhat active (7500- 10,000; n=4837)	Active ( $\geq 10,000$ ; n=2600)	Quartile 1 ( $\leq 40\%$ ; high regularity; n=6733)	Quartile 2 (40%- 51%; n=6734)	Quartile 3 (51%- 63%; n=6735)	Quartile 4 ( $>63\%$ ; low regularity; n=6733)
Systolic blood pressure (mm Hg) <sup>f,g</sup> , mean (SD)	127.4 (11.4)	128.8 (11.8)	127.2 (11.4)	125.8 (10.9)	124.5 (10.3)	126.3 (11.2)	127.2 (11.4)	127.9 (11.5)	128.1 (11.6)
<b>Systolic blood pressure (mm Hg), n (%)</b>									
≤122	7709 (28.6)	2831 (25.2)	2372 (28.7)	1571 (32.5)	935 (36.0)	2110 (31.3)	1990 (29.6)	1821 (27.0)	1788 (26.6)
122-132	7579 (28.1)	3056 (27.2)	2328 (28.2)	1404 (29.0)	791 (30.3)	1936 (28.8)	1827 (27.1)	1926 (28.6)	1890 (28.1)
>132	7646 (28.4)	3691 (32.9)	2322 (28.1)	1127 (23.3)	506 (19.5)	1698 (25.2)	1929 (28.6)	1985 (29.5)	2034 (30.1)
Unknown	4001 (14.9)	1654 (14.7)	1244 (15.0)	735 (15.2)	368 (14.2)	989 (14.7)	988 (14.7)	1003 (14.9)	1021 (15.2)
Diastolic blood pressure (mm Hg) <sup>f,g</sup> , mean (SD)	79.2 (8.2)	79.9 (8.3)	79.2 (8.2)	78.4 (8.0)	77.3 (7.8)	78.6 (8.2)	79.2 (8.2)	79.5 (8.2)	79.4 (8.2)
<b>Diastolic blood pressure (mm Hg), n (%)</b>									
≤76	7643 (28.4)	2876 (25.6)	2336 (28.3)	1504 (31.1)	927 (35.7)	2077 (30.8)	1916 (28.5)	1820 (27.0)	1830 (27.2)
76-82	7640 (28.4)	3194 (28.5)	2327 (28.2)	1379 (28.5)	740 (28.5)	1891 (28.1)	1901 (28.2)	1890 (28.1)	1958 (29.0)
>82	7651 (28.3)	3508 (31.2)	2359 (28.5)	1219 (25.2)	565 (21.6)	1776 (26.4)	1929 (28.6)	2022 (30.0)	1924 (28.6)
Unknown	4001 (14.9)	1654 (14.7)	1244 (15.0)	735 (15.2)	368 (14.2)	989 (14.7)	988 (14.7)	1003 (14.9)	1021 (15.2)
Heart rate (bpm) <sup>f,h</sup> , mean (SD)	70.6 (10.0)	71.6 (10.2)	70.4 (9.9)	69.8 (9.7)	68.6 (9.5)	71.0 (9.9)	71.0 (9.8)	70.7 (10.0)	69.8 (10.1)
<b>Heart rate (bpm), n (%)</b>									
≤66	8239 (30.6)	3071 (27.3)	2575 (31.2)	1627 (33.6)	966 (37.2)	1956 (29.1)	1962 (29.1)	2028 (30.1)	2293 (34.1)
66-75	8243 (30.6)	3409 (30.4)	2548 (30.7)	1484 (30.7)	802 (30.8)	2100 (31.2)	2054 (30.5)	2075 (30.8)	2014 (29.8)
>75	8243 (30.6)	3777 (33.6)	2477 (30.0)	1371 (28.4)	618 (23.8)	2129 (31.6)	2181 (32.4)	2083 (30.9)	1850 (27.5)
Unknown	2210 (8.2)	975 (8.7)	666 (8.1)	355 (7.3)	214 (8.2)	548 (8.1)	537 (8.0)	549 (8.2)	576 (8.6)
<b>Predominant season during the follow-up, n (%)</b>									
Winter	6153 (22.9)	2520 (22.4)	1960 (23.7)	1098 (22.7)	575 (22.0)	1559 (23.2)	1672 (24.8)	1557 (23.1)	1365 (20.3)
Spring	5739 (21.3)	2326 (20.7)	1759 (21.3)	1119 (23.1)	535 (20.6)	1593 (23.7)	1517 (22.5)	1405 (20.9)	1224 (18.2)
Summer	7627 (28.3)	3493 (31.1)	2251 (27.2)	1216 (25.2)	667 (25.7)	1559 (23.2)	1702 (25.3)	1981 (29.4)	2385 (35.4)
Fall	7416 (27.5)	2893 (25.8)	2296 (27.8)	1404 (29.0)	823 (31.7)	2022 (29.9)	1843 (27.4)	1792 (26.6)	1759 (26.1)

<sup>a</sup>M1: first 30-day period of activity tracker use.

<sup>b</sup>CV: coefficient of variation.

<sup>c</sup>Change=M1 value subtracted from the M6 value. Positive change indicates an increase during the follow-up.

<sup>d</sup>Median (IQR) are presented instead of mean (SD) due to skewed distributions.

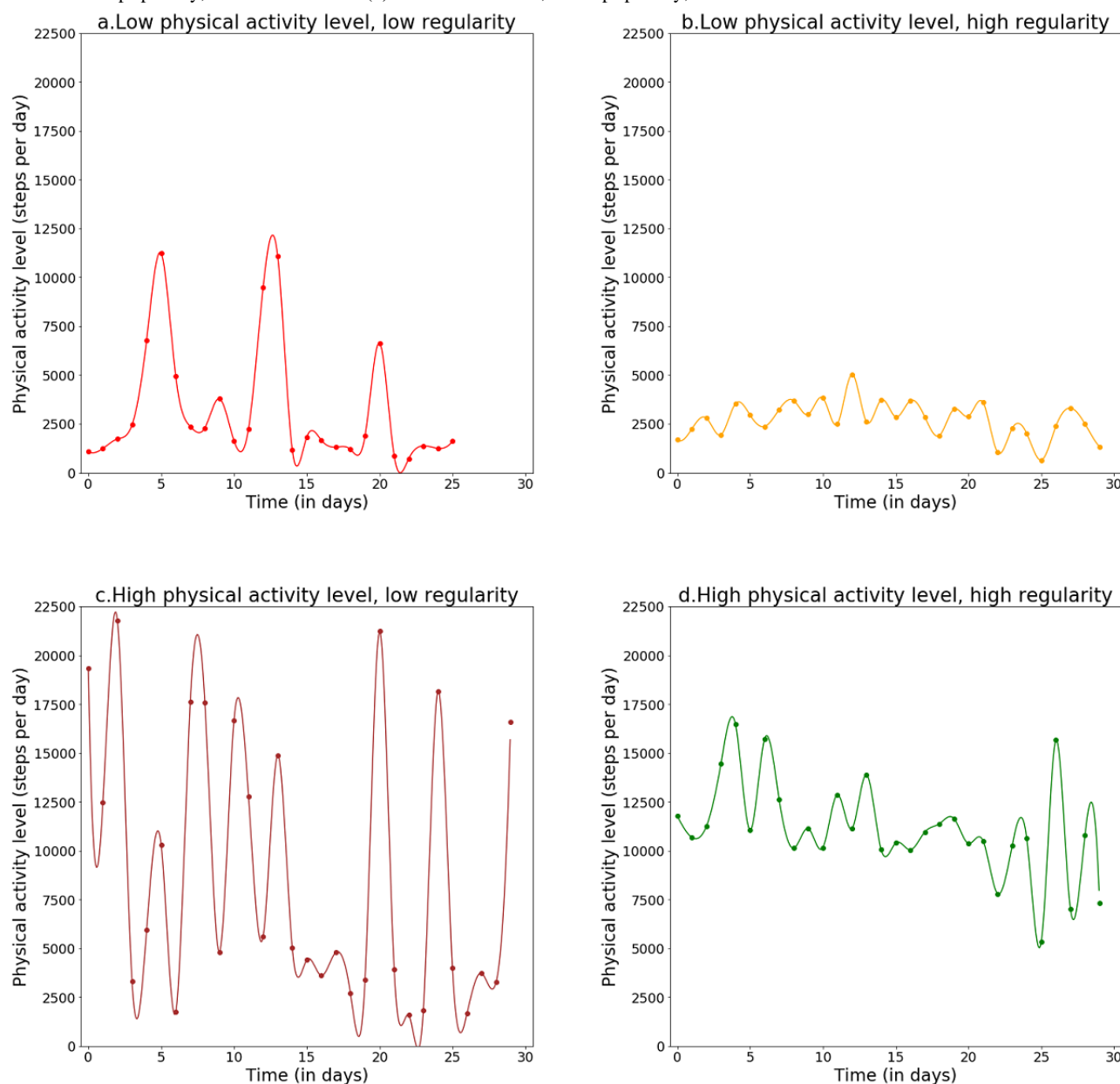
<sup>e</sup>M6: sixth 30-day period of activity tracker use.

<sup>f</sup>Evaluated using all the data available during M1 to 6 months prior.

<sup>g</sup>The mean was calculated on a sample of the study population of n=22,934 because of missing data for some participants.

<sup>h</sup>The mean was calculated on a sample of the study population of n=24,725 because of missing data for some participants.

**Figure 2.** Real-world examples of evolution of physical activity (PA) level within the first 30-day period of the use of a wearable activity tracker. (a) Mean PA level=3223 steps per day; PA level coefficient of variation (CV)=98%. (b) Mean PA level=2720 steps per day; PA level CV=34%. (c) Mean PA level=8822 steps per day; PA level CV=78%. (d) Mean PA level=11,073 steps per day; PA level CV=12%.



### PA Characteristics and 6-Month Weight Change

Associations between PA characteristics and weight changes over a 6-month period are presented in [Table 2](#).

In multivariable analyses, we found that greater mean PA levels were inversely associated with weight change. Compared with individuals who were initially in the sedentary category (<5000 steps/day), individuals who were low active (5000-7499 steps/day), somewhat active (7500-9999 steps/day), and active ( $\geq 10,000$  steps/day) had a 0.21-kg, a 0.52-kg, and a 1.17-kg greater decrease in weight (95% CI  $-0.36$  to  $-0.06$ ; 95% CI  $-0.70$  to  $-0.33$ ; and 95% CI  $-1.42$  to  $-0.93$ ), respectively. When evaluating mean PA level as a continuous variable, every 1000 steps per day increase in mean PA level was associated with a

0.11-kg decrease in weight (95% CI  $-0.14$  to  $-0.09$ ;  $P$  trend<.001).

Similarly, an inverse association between the regularity of PA level and 6-month weight change was observed ([Table 2](#)). The regression coefficients for weight change decreased progressively with increased regularity or reduced variability in the PA level at M1. Compared with users whose PA level CV was  $>63\%$ , users whose PA level CV ranged from 51% to 63% and from 40% to 51% and whose PA level CV was  $\leq 40\%$  had a 0.19-kg, a 0.23-kg, and a 0.33-kg greater decrease in weight (95% CI  $-0.38$  to  $-0.01$ ; 95% CI  $-0.41$  to  $-0.04$ ; and 95% CI  $-0.53$  to  $-0.13$ ), respectively. When evaluating PA level CV as a continuous variable, every 10% increase in PA level CV was associated with a 0.07-kg increase in weight (95% CI 0.04 to 0.11;  $P$  trend<.001).

**Table 2.** Associations of mean physical activity level and variability in physical activity level with 6-month weight change (N=26,935).

Physical activity characteristics	Model 1 <sup>a</sup>		Model 2 <sup>b</sup>	
	$\beta$ (95% CI) in kg	<i>P</i> value	$\beta$ (95% CI) in kg	<i>P</i> value
<b>Mean physical activity level<sup>c</sup></b>				
For 1000 steps/day increase <sup>d</sup>	-0.13 (-0.15 to -0.11)	<.001	-0.11 (-0.14 to -0.09)	<.001
<b>4-category step-defined physical activity<sup>e</sup></b>				
Sedentary (<5000 steps/day)	Reference	N/A <sup>f</sup>	Reference	N/A
Low active (5000-7499 steps/day)	-0.25 (-0.41 to -0.10)	.001	-0.21 (-0.36 to -0.06)	.01
Somewhat active (7500-9999 steps/day)	-0.60 (-0.78 to -0.42)	<.001	-0.52 (-0.70 to -0.33)	<.001
Active ( $\geq$ 10,000 steps/day)	-1.23 (-1.46 to -1.00)	<.001	-1.17 (-1.42 to -0.93)	<.001
<i>P</i> trend	N/A	<.001	N/A	<.001
<b>Variability in physical activity level (physical activity level coefficient of variation)<sup>e</sup></b>				
For 10% increase <sup>d</sup>	0.16 (0.12 to 0.19)	<.001	0.07 (0.04 to 0.11)	<.001
<b>Quartile groups<sup>e</sup></b>				
Q1 ( $\leq$ 40%, most regular)	-0.77 (-0.95 to -0.59)	<.001	-0.33 (-0.53 to -0.13)	.001
Q2 (40%-51%)	-0.50 (-0.68 to -0.32)	<.001	-0.23 (-0.41 to -0.04)	.02
Q3 (51%-63%)	-0.35 (-0.53 to -0.17)	<.001	-0.19 (-0.38 to -0.01)	.04
Q4 (>63%, less regular)	Reference	N/A	Reference	N/A
<i>P</i> trend	N/A	<.001	N/A	<.001

<sup>a</sup>Model 1: univariable.

<sup>b</sup>Model 2: adjusted for age, sex, systolic blood pressure, diastolic blood pressure, heart rate, and for the predominant season during the follow-up.

<sup>c</sup>Evaluated during the first 30-day period of activity tracker use.

<sup>d</sup>Continuous variables of mean physical activity level and physical activity level coefficient of variation were simultaneously included in the multivariable models.

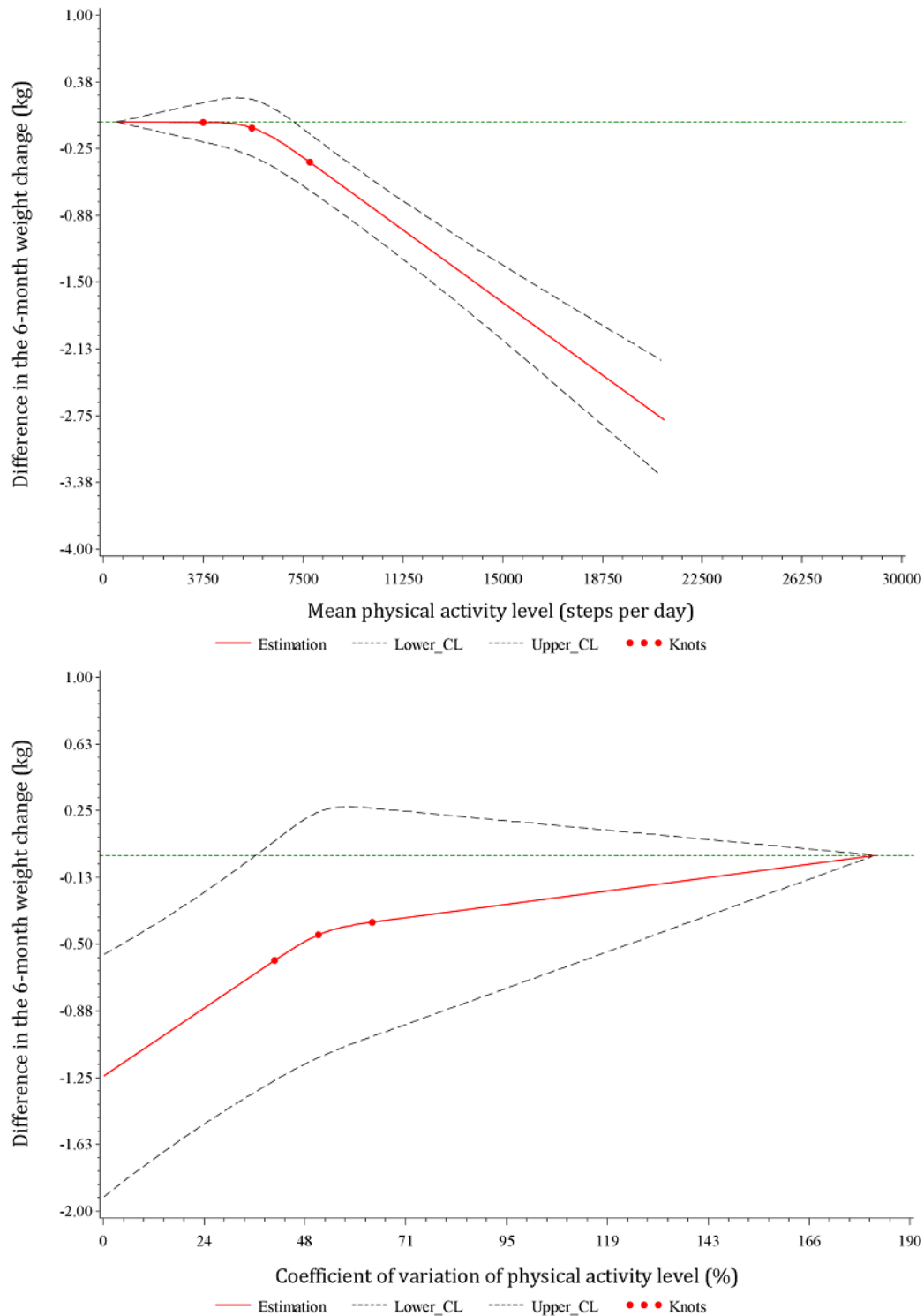
<sup>e</sup>Categorical variables of mean physical activity level and physical activity level coefficient of variation were simultaneously included in the multivariable models.

<sup>f</sup>N/A: not applicable.

The restricted cubic spline regression analysis specified the continuous relationships between the mean PA level and regularity and the 6-month weight change (Figure 3). Figure 3 (upper part) shows a dose-response association between mean PA level and 6-month weight change, where 516 steps per day was the reference value for the mean PA level. A plateau was

observed between the reference value and approximately 7500 steps per day, followed by a decrease in weight change for individuals with a mean PA level above 7500 steps per day. In Figure 3 (lower part), we observed a trend toward weight loss with an increase in regularity in PA level.

**Figure 3.** Relationships between mean physical activity level and the variability in physical activity level and 6-month weight change fitted with restricted cubic splines (3 knots placed at the 25th, 50th, and 75th percentiles). The solid red line represents the mean differences; the dashed lines are the 95% confidence limits; the green straight horizontal line corresponds to the line,  $y=0$ . The y-axis represents the difference in the 6-month weight change between individuals with any value of mean physical activity level (top) or physical activity level coefficient of variation (bottom) with individuals with the minimum value of mean physical activity level (516 steps/day; top) or with the maximum value of physical activity level coefficient of variation (182%; bottom). CL: confidence limit.



**Secondary Analysis: 6-Month Change in PA Characteristics and 6-Month Weight Change**

Associations between 6-month change in PA characteristics and weight change are presented in [Table 3](#).

An inverse association was observed between the 6-month change in mean PA level and the 6-month weight change. A 1000 steps per day increase in change in mean PA level resulted in a 0.26-kg decrease in weight (95% CI -0.29 to -0.23), whereas an increase from one category to another in the



4-category step-defined PA classification was associated with a 0.54-kg decrease in weight (95% CI -0.61 to -0.46). Compared with individuals whose change in mean PA level ranged from -1000 to +1000 steps per day, individuals whose change in mean PA level ranged from -1000 to +3000 steps per day and individuals whose change was above 3000 steps

per day had a 0.68-kg and a 2.35-kg greater decrease in weight (95% CI -0.85 to -0.51 and 95% CI -2.57 to -2.12), respectively.

No association was found between the 6-month change in the variability of PA level and weight change.

**Table 3.** Associations of 6-month change in physical activity level and variability with 6-month weight change (N=26,935).

Physical activity characteristics	Model 1 <sup>a</sup>		Model 2 <sup>b</sup>	
	$\beta$ (95% CI) in kg	P value	$\beta$ (95% CI) in kg	P value
<b>6-month change in mean physical activity level<sup>c</sup></b>				
For 1000 steps/day increase	-0.28 (-0.31 to -0.25)	<.001	-0.26 (-0.29 to -0.23)	<.001
<b>5-category change in mean physical activity level</b>				
Decrease of more than 3000 steps/day	-0.02 (-0.28 to 0.23)	.87	-0.10 (-0.36 to 0.15)	.43
Decrease between 3000 and 1000 steps/day	0.06 (-0.11 to 0.24)	.46	0.003 (-0.168 to 0.174)	.97
Change between -1000 and +1000 steps/day	Reference	N/A <sup>d</sup>	Reference	N/A
Increase between 1000 and 3000 steps/day	-0.73 (-0.90 to -0.56)	<.001	-0.68 (-0.85 to -0.51)	<.001
Increase of more than 3000 steps/day	-2.42 (-2.65 to -2.19)	<.001	-2.35 (-2.57 to -2.12)	<.001
Change in 4-category step-defined physical activity	-0.59 (-0.66 to -0.52)	<.001	-0.54 (-0.61 to -0.46)	<.001
<b>6-month change in variability of physical activity level (change in physical activity level coefficient of variation)<sup>c</sup></b>				
For 10% increase	0.02 (-0.02 to 0.05)	.34	0.02 (-0.01 to 0.05)	.15
Change in 4-category physical activity level CV	0.03 (-0.02 to 0.09)	.21	0.04 (-0.01 to 0.10)	.09

<sup>a</sup>Model 1: univariable.

<sup>b</sup>Model 2: adjusted for age, sex, systolic blood pressure, diastolic blood pressure, heart rate, and for the predominant season during the follow-up.

<sup>c</sup>Change=M1 value subtracted from the M6 value. Positive change indicates an increase during the follow-up.

<sup>d</sup>N/A: not applicable.

## Discussion

### Principal Findings and Comparison With the Literature

On the basis of the data from more than 25,000 connected device users, we were able to study, in free-living conditions, the associations between PA characteristics (mean level and variability), their 6-month changes, and the 6-month weight change.

Our results suggest that greater mean PA levels are associated with weight loss. Taking an additional 1000 steps per day was associated with a 0.11-kg greater weight loss over a 6-month period. In contrast with our findings, Unick et al [17] found no association between total daily steps and weight change among individuals with normal weight or overweight after 4 months, 1 year, or 2 years of a weight gain prevention intervention. However, they worked on a smaller sample of 599 participants who were predominantly female, and they compared daily steps among participants who gained >1 lb (approximately 0.45 kg) to those who lost weight or gained  $\leq$ 1 lb. These differences may explain the discrepancies in the findings. Furthermore, the choice of a 1-lb cut-off point to compare participants gaining weight to those who lost weight may be too low for a 2-year intervention. In addition, Thomson et al [27] did not find any association between initial PA level (baseline mean steps/day)

and 6-month weight change in a walking intervention study targeting a reduction in blood pressure. This study involved a smaller sample (179 adults) with a potentially lower variability in the studied PA profiles, which could explain the difference with our results.

When we analyzed the initial mean PA levels expressed according to the 4-category step-defined PA classification, we observed that PA levels between 5000 and 7499 steps per day, between 7500 and 9999 steps per day, and above  $\geq$ 10,000 steps per day were all associated with a greater reduction in weight after 6 months in comparison with mean PA levels <5000 steps per day, which is considered as a threshold for inactivity [13,32]. Higher categories were more strongly associated with weight loss than lower categories, suggesting that “some PA is good, more PA is better” with respect to weight change. This result is in line with the study of McTiernan et al [33] that showed that higher exercise levels were linked to greater weight loss. This finding seemed logical from a purely physiological and physical point of view, that is, the greater the PA, the greater the energy expenditure, and the greater the weight loss if constant energy intakes are maintained. Our restricted cubic spline regression analysis, which was performed to better characterize the dose-response continuous relationship between mean PA level and the 6-month weight change, supported this finding. This suggests that mean PA levels above approximately 7500 steps per day were significantly associated with greater

weight loss compared with the minimum value of 516 steps per day (Figure 3). Tudor-Locke et al [13] revisited the 10,000 steps per day cut-off point and proposed step-based recommendations congruent with current public health PA guidelines. These recommendations suggest that the total daily volume of PA associated with meeting the minimum recommended level is 7000 to 8000 steps per day [13,34]. Thus, our study suggests that having a PA level consistent with the current public health guidelines can be associated with weight loss. A recently published systematic review reported a dose-response relationship between daily step counts and mortality [35]. In our study, a greater decrease in weight was observed even at low levels of PA, below the commonly proposed threshold of 10,000 steps per day. Hall et al [35] showed that a reduced risk of mortality was also observed below the threshold of 10,000 steps per day.

For changes in PA levels, we observed that each 1000 steps per day increase in PA level over the 6-month follow-up was associated with a 0.26-kg greater decrease in body weight. This finding is in line with those of previous studies. Thomson et al [27] found a similar association between a 6-month change in mean steps per day and BMI change. They reported that a 1000 steps per day increase in change in PA level resulted in a 0.13-kg/m<sup>2</sup> greater decrease in BMI, which corresponds to an approximately 0.38-kg greater reduction in weight considering an average height of 1.70 m [27]. This slightly stronger magnitude in the relationship could be explained by the higher initial mean PA level and mean change in PA level within the population in the Thomson et al study [27] (7279 steps/day, SD 3417 vs 5940 steps/day, SD 2930 for initial PA level and 1855 steps/day, SD 2710 vs 144 steps/day, SD 2306 for change in PA level). However, in our study, changes in PA level were evaluated by subtracting the M1 mean daily steps from the M6 mean daily steps, which is more accurate and reflects the 6-month absolute change. In contrast, in the Thomson et al study [27], the concept of change in PA was vaguer given that changes in PA corresponded to differences between steps per day values assessed during the first 2 weeks of the study and those evaluated during the remaining weeks (from the third to the 27th week). Creasy et al [36] also reported a significant association between PA level and weight change across an 18-month intervention study, suggesting an additional 0.21-kg of weight loss for each additional 1000 steps per day. Similarly, we noticed that a one-category change in the 4-category step-defined PA classification was associated with a 0.54-kg greater decrease in weight. Ganesan et al [37] found that a one-category elevation was associated with a weight decrease of 0.23 kg over a 3-month period. Similar step count categories to the 4-category step-defined PA classification were used in their study. We also observed a dose-response relationship between changes in PA level and weight change, with higher increases over time in PA level linked to greater weight loss.

In this study, we also observed that lower variability in PA levels, or greater regularity in PA levels, was significantly associated with greater weight loss. However, no association was found between the 6-month change in the variability of PA level and weight change. Even though changes in variability were not linked to weight change, the fact that the initial

variability in PA level was associated with weight loss regardless of PA level suggests that step count variability can be a relevant PA component that deserves further exploration in relation to weight change and other health parameters. In addition, our results highlight the complementarity between the mean PA level and variability in PA level. Figure 2 reflects the idea that individuals with similar mean PA levels can have very different variabilities. The notion of variability in health-related indicators or behaviors has already been studied for many years in the case of glycaemia [38-41] and more recently for sleep duration [42]. These studies showed that variability in health indicators or behaviors could have harmful effects on health. This should prompt us to extend the research on variability to other health behaviors, including PA.

### Strengths and Limitations

This study has numerous strengths, including a large study population size. Indeed, it is the first of its kind to concomitantly study both PA level and variability in PA level, their changes, and weight change over a 6-month period. It is further strengthened by the fact that PA was assessed objectively via wearable activity trackers used by consumers in real-life settings and not through self-reported measures that are known to be prone to social desirability and recall bias [17]. Furthermore, most of the studies in the literature used only 3 to 7 daily step measures to assess PA of their participants during a given period [17,32,36], whereas our study population turned out to be composed of frequent users, given that 75% of the users had at least 20 daily step measures during M1 and M6. This enabled us to analyze how the variability in PA level, an indicator poorly studied in the literature, was related to weight change. Moreover, PA level was studied using the 1000 steps per day increment and the recognized 4-category step-defined PA classification proposed by Tudor-Locke et al [12,13], allowing for comparison and fostering harmonization across studies.

This study had some limitations. First, the study sample was mainly composed of men; a similar study should be conducted in another population with a larger proportion of women. The study population had a relatively high weight value at baseline (88.9, SD 18.9 kg), which may limit the generalizability of the findings to the general population. It is also possible that our findings would not be generalizable to the general population given that the individuals included in this study, consumers who chose to buy a Withings wearable activity monitor and a weighing scale, may have been more motivated to lose weight. However, we believe that any additional motivation for weight loss would be comparable with that of other populations analyzed in similar studies, such as weight loss intervention programs. In addition, the mean PA level in our study population (5940 steps/day, SD 2930) was similar to that of samples from other studies, including the Women's Health Study cohort (5499 steps/day [34]) and the NAVIGATOR (Nateglinide and Valsartan in Impaired Glucose Tolerance Outcomes Research) study (6205 steps/day, SD 3727 [43]), showing that individuals included in our study were not particularly more active than other study populations. Information on the height of users was not available, preventing us from calculating the BMI and consequently from studying individuals who are overweight or obese specifically or from investigating if the associations

observed vary by weight status. Nevertheless, we adjusted our analysis for several health-related parameters strongly correlated with BMI, such as blood pressure and heart rate. Information on food intake was not available, so we could neither control for it in our analyses nor assess the potential relative contribution of dietary intake to weight change. Finally, the intensities of the PAs performed were not accessible, as we only had access to daily step data. Nonetheless, daily step count is an increasingly widespread and intuitive metric that can be easily used as a target for health benefits. Indeed, with the expansion of wearable activity monitors and smartphones in the commercial market, the daily step count metric has become largely accessible among the general population, which justifies its use and emphasizes the need to study it in relation to health outcomes [35]. Working with data based on the number of steps accumulated during the day also implies that we only have information on ambulatory PAs [32]. Thus, participation in nonambulatory PAs, such as swimming or cycling, may not have been taken into account and properly considered. However, it is known that ambulatory PAs such as walking remain a central component of PA in the general population [35].

### Perspectives and Conclusions

Further investigations with complementary data on diet, sedentary behaviors, and PA intensity or cadence (steps/min), which will allow the calculation of time spent in moderate-to-vigorous and light PA intensities, are warranted. Alternative and/or more innovative methods such as unsupervised clustering methods (eg, k-means or hierarchical agglomerative clustering) or latent class mixed models could be used on data collected from consumer wearable devices to identify temporal evolution patterns or groups of individuals with similar behaviors and study them in relation to various health outcomes. Indeed, we also encourage our study to be replicated by focusing on other health-related parameters different from weight change such as blood pressure, heart rate, or fat mass percentage [44]. These are easily measurable with the use of connected devices and all represent key cardiometabolic risk factors to study in relation to device-assessed PA within the context of obesity epidemic and the digital revolution [44]. This study has also shown that connected devices can serve as useful tools to track PA and weight in large epidemiological studies. Future weight loss

intervention studies are highly encouraged to use connected activity trackers to objectively and accurately monitor the physical behaviors of their participants. The use of activity trackers, which is increasingly becoming mainstream, should now be leveraged to track and enhance PA in order to derive truly personalized prevention programs, adapted to the individual's lifestyle and living conditions and to move toward personalized prevention. Machine learning approaches, including neural networks, can be explored to predict, for instance, the days when a user would be less active. Such information could be used by health-related apps to send motivational messages or propose services to the users to help them become more active. More generally, we strongly promote the creation of collaborations between academic researchers and wearable device manufacturers, which could give researchers access to rich and valuable databases. Future studies on these sources of data could help improve epidemiological and health knowledge, as well as enhance health-related wearable devices.

From a public health perspective, our results suggest that weight-related health benefits can be observed below the controversial 10,000 steps per day threshold and emphasize the idea that "some PA is good, more PA is better" with regard to weight loss [11,32,35]. Our work may have important public health implications when encouraging adults to engage in PA that is monitored as steps per day, especially adults who are low active for whom adherence to the 10,000 steps per day may be too ambitious or unrealistic [34].

In conclusion, greater baseline daily step counts were associated with a greater decrease in weight. More importantly, our findings also indicate that increasing PA levels over time, irrespective of the baseline level, may be beneficial in the short term. In addition, a more regular level of PA should be promoted. Indeed, our results suggest that the variability in PA level is an interesting additional parameter to be considered as a digital biomarker candidate when assessing the impact of PA on health, which deserves to be further considered and studied in weight loss programs and observational studies. This work has shown that data from wearable devices are helpful for the digital phenotyping of large populations in real-life settings and also to suggest new metrics to characterize PA that could be easily replicated in future studies.

### Acknowledgments

The authors would like to thank the anonymous reviewers for their thorough review and Dr Joseph Rothwell for careful proofreading of the manuscript.

DE was supported by a doctoral grant from the French National Cancer Institute (INCa\_13563). This study was also supported by the French Research Agency ANR (Agence Nationale de la Recherche) via an "Investissement d'Avenir" grant (investment for the future grant, ANR-10-COHO-0006), which supports the E4N study. GF is supported by the Luxembourg Institute of Health. CG is supported by a doctoral grant from LIONS, a Luxembourgish association that supports cancer research. Study funders and sponsors had no role in the design of the study, analysis or interpretation of data, writing of the manuscript, or decision to submit the manuscript for publication.

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request and with permission from the Withings company.

## Authors' Contributions

DE and GF formulated the research questions and conceived and designed the study. DE performed the statistical analysis and drafted the initial manuscript. All authors contributed to the interpretation of data, discussed the manuscript, reviewed the paper, revised it critically, and approved its final version to be published. GF is the guarantor of this work and, as such, has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## Conflicts of Interest

DC is employed by Withings. All other authors have no conflicts to declare.

## References

1. Obesity and overweight. World Health Organization. 2020. URL: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> [accessed 2019-09-23]
2. The National Heart, Lung, and Blood Institute Expert Panel on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. Executive Summary of the Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. *Journal of the American Nutrition and Dietetics* 1998 Oct;98(10):1178-1191. [doi: [10.1016/s0002-8223\(98\)00276-4](https://doi.org/10.1016/s0002-8223(98)00276-4)]
3. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *The Lancet* 2014 Aug;384(9945):755-765. [doi: [10.1016/s0140-6736\(14\)60892-8](https://doi.org/10.1016/s0140-6736(14)60892-8)]
4. Preface to the Expert Panel Report (comprehensive version which includes systematic evidence review, evidence statements, and recommendations). *Obesity (Silver Spring)* 2014 Jul;22 Suppl 2:S40 [FREE Full text] [doi: [10.1002/oby.20822](https://doi.org/10.1002/oby.20822)] [Medline: [24961824](https://pubmed.ncbi.nlm.nih.gov/24961824/)]
5. Magkos F, Fraterrigo G, Yoshino J, Luecking C, Kirbach K, Kelly S, et al. Effects of Moderate and Subsequent Progressive Weight Loss on Metabolic Function and Adipose Tissue Biology in Humans with Obesity. *Cell Metab* 2016 Apr 12;23(4):591-601 [FREE Full text] [doi: [10.1016/j.cmet.2016.02.005](https://doi.org/10.1016/j.cmet.2016.02.005)] [Medline: [26916363](https://pubmed.ncbi.nlm.nih.gov/26916363/)]
6. Colberg SR, Sigal RJ, Yardley JE, Riddell MC, Dunstan DW, Dempsey PC, et al. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. *Diabetes Care* 2016 Nov 25;39(11):2065-2079 [FREE Full text] [doi: [10.2337/dc16-1728](https://doi.org/10.2337/dc16-1728)] [Medline: [27926890](https://pubmed.ncbi.nlm.nih.gov/27926890/)]
7. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE. Beneficial effects of exercise: shifting the focus from body weight to other markers of health. *Br J Sports Med* 2009 Dec 29;43(12):924-927. [doi: [10.1136/bjsm.2009.065557](https://doi.org/10.1136/bjsm.2009.065557)] [Medline: [19793728](https://pubmed.ncbi.nlm.nih.gov/19793728/)]
8. Alevizos A, Lentzas J, Kokkoris S, Mariolis A, Korantzopoulos P. Physical activity and stroke risk. *Int J Clin Pract* 2005 Aug;59(8):922-930. [doi: [10.1111/j.1742-1241.2005.00536.x](https://doi.org/10.1111/j.1742-1241.2005.00536.x)] [Medline: [16033614](https://pubmed.ncbi.nlm.nih.gov/16033614/)]
9. Miller TD, Balady GJ, Fletcher GF. Exercise and its role in the prevention and rehabilitation of cardiovascular disease. *ann. behav. med* 1997 Sep;19(3):220-229. [doi: [10.1007/bf02892287](https://doi.org/10.1007/bf02892287)]
10. Caspersen CJ, Powell K, Christenson G. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100(2):126-131 [FREE Full text] [Medline: [3920711](https://pubmed.ncbi.nlm.nih.gov/3920711/)]
11. Haseler C, Crooke R, Haseler T. Promoting physical activity to patients. *BMJ* 2019 Sep 17;366:l5230. [doi: [10.1136/bmj.l5230](https://doi.org/10.1136/bmj.l5230)] [Medline: [31530549](https://pubmed.ncbi.nlm.nih.gov/31530549/)]
12. Tudor-Locke C, Bassett DR. How many steps/day are enough? Preliminary pedometer indices for public health. *Sports Med* 2004;34(1):1-8. [doi: [10.2165/00007256-200434010-00001](https://doi.org/10.2165/00007256-200434010-00001)] [Medline: [14715035](https://pubmed.ncbi.nlm.nih.gov/14715035/)]
13. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De Cocker K, Giles-Corti B, et al. How many steps/day are enough? For adults. *Int J Behav Nutr Phys Act* 2011 Jul 28;8(1):79 [FREE Full text] [doi: [10.1186/1479-5868-8-79](https://doi.org/10.1186/1479-5868-8-79)] [Medline: [21798015](https://pubmed.ncbi.nlm.nih.gov/21798015/)]
14. Swift DL, Johannsen NM, Lavie CJ, Earnest CP, Church TS. The role of exercise and physical activity in weight loss and maintenance. *Prog Cardiovasc Dis* 2014 Jan;56(4):441-447 [FREE Full text] [doi: [10.1016/j.pcad.2013.09.012](https://doi.org/10.1016/j.pcad.2013.09.012)] [Medline: [24438736](https://pubmed.ncbi.nlm.nih.gov/24438736/)]
15. Jakicic JM. The effect of physical activity on body weight. *Obesity (Silver Spring)* 2009 Dec;17 Suppl 3:S34-S38 [FREE Full text] [doi: [10.1038/oby.2009.386](https://doi.org/10.1038/oby.2009.386)] [Medline: [19927144](https://pubmed.ncbi.nlm.nih.gov/19927144/)]
16. Catenacci VA, Wyatt HR. The role of physical activity in producing and maintaining weight loss. *Nat Clin Pract Endocrinol Metab* 2007 Jul;3(7):518-529 [FREE Full text] [doi: [10.1038/ncpendmet0554](https://doi.org/10.1038/ncpendmet0554)] [Medline: [17581621](https://pubmed.ncbi.nlm.nih.gov/17581621/)]
17. Unick JL, Lang W, Williams SE, Bond DS, Egan CM, Espeland MA, SNAP Research Group. Objectively-assessed physical activity and weight change in young adults: a randomized controlled trial. *Int J Behav Nutr Phys Act* 2017 Dec 04;14(1):165 [FREE Full text] [doi: [10.1186/s12966-017-0620-x](https://doi.org/10.1186/s12966-017-0620-x)] [Medline: [29202850](https://pubmed.ncbi.nlm.nih.gov/29202850/)]
18. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport* 2000 Jun 13;71 Suppl 2(sup2):1-14. [doi: [10.1080/02701367.2000.11082780](https://doi.org/10.1080/02701367.2000.11082780)] [Medline: [25680007](https://pubmed.ncbi.nlm.nih.gov/25680007/)]
19. Cox CE. Role of Physical Activity for Weight Loss and Weight Maintenance. *Diabetes Spectr* 2017 Aug 15;30(3):157-160 [FREE Full text] [doi: [10.2337/ds17-0013](https://doi.org/10.2337/ds17-0013)] [Medline: [28848307](https://pubmed.ncbi.nlm.nih.gov/28848307/)]

20. 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](https://doi.org/10.2196/11819)] [Medline: [30977740](https://pubmed.ncbi.nlm.nih.gov/30977740/)]
21. Pourzanjani A, Quisel T, Foschini L. Adherent Use of Digital Health Trackers Is Associated with Weight Loss. *PLoS One* 2016 Apr 6;11(4):e0152504 [FREE Full text] [doi: [10.1371/journal.pone.0152504](https://doi.org/10.1371/journal.pone.0152504)] [Medline: [27049859](https://pubmed.ncbi.nlm.nih.gov/27049859/)]
22. Painter SL, Ahmed R, Hill JO, Kushner RF, Lindquist R, Brunning S, et al. What Matters in Weight Loss? An In-Depth Analysis of Self-Monitoring. *J Med Internet Res* 2017 May 12;19(5):e160 [FREE Full text] [doi: [10.2196/jmir.7457](https://doi.org/10.2196/jmir.7457)] [Medline: [28500022](https://pubmed.ncbi.nlm.nih.gov/28500022/)]
23. 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 Mar 01;312(3):R358-R367 [FREE Full text] [doi: [10.1152/ajpregu.00349.2016](https://doi.org/10.1152/ajpregu.00349.2016)] [Medline: [28052867](https://pubmed.ncbi.nlm.nih.gov/28052867/)]
24. Kaewkannate K, Kim S. A comparison of wearable fitness devices. *BMC Public Health* 2016 May 24;16(1):433 [FREE Full text] [doi: [10.1186/s12889-016-3059-0](https://doi.org/10.1186/s12889-016-3059-0)] [Medline: [27220855](https://pubmed.ncbi.nlm.nih.gov/27220855/)]
25. Sperrin M, Rushton H, Dixon WG, Normand A, Villard J, Chieh A, et al. Who Self-Weighs and What Do They Gain From It? A Retrospective Comparison Between Smart Scale Users and the General Population in England. *J Med Internet Res* 2016 Jan 21;18(1):e17 [FREE Full text] [doi: [10.2196/jmir.4767](https://doi.org/10.2196/jmir.4767)] [Medline: [26794900](https://pubmed.ncbi.nlm.nih.gov/26794900/)]
26. Hay A, Ayis S, Nzelu D, James L, Kametas NA. Validation of the Withings BP-800 in pregnancy and impact of maternal characteristics on the accuracy of blood pressure measurement. *Pregnancy Hypertens* 2016 Oct;6(4):406-412. [doi: [10.1016/j.preghy.2016.09.005](https://doi.org/10.1016/j.preghy.2016.09.005)] [Medline: [27939491](https://pubmed.ncbi.nlm.nih.gov/27939491/)]
27. Thomson JL, Landry AS, Zoellner JM, Tudor-Locke C, Webster M, Connell C, et al. Several steps/day indicators predict changes in anthropometric outcomes: HUB City Steps. *BMC Public Health* 2012 Nov 15;12(1):983 [FREE Full text] [doi: [10.1186/1471-2458-12-983](https://doi.org/10.1186/1471-2458-12-983)] [Medline: [23153060](https://pubmed.ncbi.nlm.nih.gov/23153060/)]
28. Menai M, Brouard B, Vegreville M, Chieh A, Schmidt N, Oppert J, et al. Cross-Sectional and longitudinal associations of objectively-measured physical activity on blood pressure: evaluation in 37 countries. *Health Promot Perspect* 2017 Sep 26;7(4):190-196 [FREE Full text] [doi: [10.15171/hpp.2017.34](https://doi.org/10.15171/hpp.2017.34)] [Medline: [29085795](https://pubmed.ncbi.nlm.nih.gov/29085795/)]
29. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act* 2015 Mar 27;12(1):42 [FREE Full text] [doi: [10.1186/s12966-015-0201-9](https://doi.org/10.1186/s12966-015-0201-9)] [Medline: [25890168](https://pubmed.ncbi.nlm.nih.gov/25890168/)]
30. An H, Jones GC, Kang S, Welk GJ, Lee J. How valid are wearable physical activity trackers for measuring steps? *Eur J Sport Sci* 2017 Apr 02;17(3):360-368. [doi: [10.1080/17461391.2016.1255261](https://doi.org/10.1080/17461391.2016.1255261)] [Medline: [27912681](https://pubmed.ncbi.nlm.nih.gov/27912681/)]
31. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci Med Rehabil* 2015 Oct 12;7(1):24 [FREE Full text] [doi: [10.1186/s13102-015-0018-5](https://doi.org/10.1186/s13102-015-0018-5)] [Medline: [26464801](https://pubmed.ncbi.nlm.nih.gov/26464801/)]
32. Pillay JD, van der Ploeg HP, Kolbe-Alexander TL, Proper KI, van Stralen M, Tomaz SA, et al. The association between daily steps and health, and the mediating role of body composition: a pedometer-based, cross-sectional study in an employed South African population. *BMC Public Health* 2015 Feb 22;15(1):174 [FREE Full text] [doi: [10.1186/s12889-015-1381-6](https://doi.org/10.1186/s12889-015-1381-6)] [Medline: [25885183](https://pubmed.ncbi.nlm.nih.gov/25885183/)]
33. McTiernan A, Sorensen B, Irwin ML, Morgan A, Yasui Y, Rudolph RE, et al. Exercise effect on weight and body fat in men and women. *Obesity (Silver Spring)* 2007 Jun;15(6):1496-1512 [FREE Full text] [doi: [10.1038/oby.2007.178](https://doi.org/10.1038/oby.2007.178)] [Medline: [17557987](https://pubmed.ncbi.nlm.nih.gov/17557987/)]
34. Lee I, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of Step Volume and Intensity With All-Cause Mortality in Older Women. *JAMA Intern Med* 2019 May 29;179(8):1105 [FREE Full text] [doi: [10.1001/jamainternmed.2019.0899](https://doi.org/10.1001/jamainternmed.2019.0899)] [Medline: [31141585](https://pubmed.ncbi.nlm.nih.gov/31141585/)]
35. Hall KS, Hyde ET, Bassett DR, Carlson SA, Carnethon MR, Ekelund U, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. *Int J Behav Nutr Phys Act* 2020 Jun 20;17(1):78 [FREE Full text] [doi: [10.1186/s12966-020-00978-9](https://doi.org/10.1186/s12966-020-00978-9)] [Medline: [32563261](https://pubmed.ncbi.nlm.nih.gov/32563261/)]
36. Creasy SA, Lang W, Tate DF, Davis KK, Jakicic JM. Pattern of Daily Steps is Associated with Weight Loss: Secondary Analysis from the Step-Up Randomized Trial. *Obesity (Silver Spring)* 2018 Jun 06;26(6):977-984 [FREE Full text] [doi: [10.1002/oby.22171](https://doi.org/10.1002/oby.22171)] [Medline: [29633583](https://pubmed.ncbi.nlm.nih.gov/29633583/)]
37. 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](https://doi.org/10.1016/j.jacc.2016.03.472)] [Medline: [27050185](https://pubmed.ncbi.nlm.nih.gov/27050185/)]
38. Wang A, Liu X, Xu J, Han X, Su Z, Chen S, et al. Visit-to-Visit Variability of Fasting Plasma Glucose and the Risk of Cardiovascular Disease and All-Cause Mortality in the General Population. *JAHA* 2017 Dec 02;6(12):e006757. [doi: [10.1161/jaha.117.006757](https://doi.org/10.1161/jaha.117.006757)]
39. Yu JH, Han K, Park S, Lee DY, Nam GE, Seo JA, et al. Effects of long-term glycemic variability on incident cardiovascular disease and mortality in subjects without diabetes: A nationwide population-based study. *Medicine (Baltimore)* 2019 Jul;98(29):e16317 [FREE Full text] [doi: [10.1097/MD.00000000000016317](https://doi.org/10.1097/MD.00000000000016317)] [Medline: [31335679](https://pubmed.ncbi.nlm.nih.gov/31335679/)]

40. Jang J, Moon S, Cho S, Cho KH, Oh C. Visit-to-visit HbA1c and glucose variability and the risks of macrovascular and microvascular events in the general population. *Sci Rep* 2019 Feb 04;9(1):1374 [FREE Full text] [doi: [10.1038/s41598-018-37834-7](https://doi.org/10.1038/s41598-018-37834-7)] [Medline: [30718626](https://pubmed.ncbi.nlm.nih.gov/30718626/)]
41. Kobayashi D, Noto H, Takahashi O, Shimbo T. Glycemic variability and subsequent malignancies among the population without diabetes. *Diabetes Res Clin Pract* 2020 Jan;159:107987. [doi: [10.1016/j.diabres.2019.107987](https://doi.org/10.1016/j.diabres.2019.107987)] [Medline: [31866528](https://pubmed.ncbi.nlm.nih.gov/31866528/)]
42. Huang T, Mariani S, Redline S. Sleep Irregularity and Risk of Cardiovascular Events: The Multi-Ethnic Study of Atherosclerosis. *J Am Coll Cardiol* 2020 Mar 10;75(9):991-999 [FREE Full text] [doi: [10.1016/j.jacc.2019.12.054](https://doi.org/10.1016/j.jacc.2019.12.054)] [Medline: [32138974](https://pubmed.ncbi.nlm.nih.gov/32138974/)]
43. Kraus WE, Yates T, Tuomilehto J, Sun J, Thomas L, McMurray JJV, et al. Relationship between baseline physical activity assessed by pedometer count and new-onset diabetes in the NAVIGATOR trial. *BMJ Open Diabetes Res Care* 2018 Jul 19;6(1):e000523. [doi: [10.1136/bmjdr-2018-000523](https://doi.org/10.1136/bmjdr-2018-000523)] [Medline: [30073088](https://pubmed.ncbi.nlm.nih.gov/30073088/)]
44. 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 Jun 3;2(1):45 [FREE Full text] [doi: [10.1038/s41746-019-0121-1](https://doi.org/10.1038/s41746-019-0121-1)] [Medline: [31304391](https://pubmed.ncbi.nlm.nih.gov/31304391/)]

## Abbreviations

**CV:** coefficient of variation

**PA:** physical activity

*Edited by L Buis; submitted 30.10.20; peer-reviewed by C Forbes, S Rostam Niakan Kalhori, A Kingsnorth, U Reimer; comments to author 11.12.20; revised version received 31.12.20; accepted 26.02.21; published 15.04.21.*

*Please cite as:*

*El Fatouhi D, Delrieu L, Goetzinger C, Malisoux L, Affret A, Campo D, Fagherazzi G*

*Associations of Physical Activity Level and Variability With 6-Month Weight Change Among 26,935 Users of Connected Devices: Observational Real-Life Study*

*JMIR Mhealth Uhealth* 2021;9(4):e25385

URL: <https://mhealth.jmir.org/2021/4/e25385>

doi: [10.2196/25385](https://doi.org/10.2196/25385)

PMID: [33856352](https://pubmed.ncbi.nlm.nih.gov/33856352/)

©Douae El Fatouhi, Lidia Delrieu, Catherine Goetzinger, Laurent Malisoux, Aurélie Affret, David Campo, Guy Fagherazzi. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 15.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Corrigenda and Addenda

# Correction: A Novel Remote Follow-Up Tool Based on an Instant Messaging/Social Media App for the Management of Patients With Low Anterior Resection Syndrome: Pilot Prospective Self-Control Study

Fan Liu<sup>1,2\*</sup>, PhD, MD; Peng Guo<sup>1,2\*</sup>, PhD, MD; Xiangqian Su<sup>3</sup>, PhD, MD; Ming Cui<sup>3</sup>, PhD, MD; Jianlong Jiang<sup>4</sup>, MD; Suo Wang<sup>4</sup>, MSc, MD; Zhouman Yu<sup>5</sup>, MD; Runhe Zhou<sup>5</sup>, MD; Yingjiang Ye<sup>1,2</sup>, PhD, MD

<sup>1</sup>Department of Gastroenterological Surgery, Peking University People's Hospital, Beijing, China

<sup>2</sup>Beijing Key Laboratory of Colorectal Cancer Diagnosis and Treatment Research, Beijing, China

<sup>3</sup>Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Gastrointestinal Surgery IV, Peking University Cancer Hospital and Institute, Beijing, China

<sup>4</sup>Department of General Surgery, Changshu Hospital Affiliated to Soochow University, First People's Hospital of Changshu City, Changshu, China

<sup>5</sup>Department of General Surgery, QiLu Hospital (Qingdao), Cheeloo College of Medicine, Shandong University, Qingdao, China

\* these authors contributed equally

**Corresponding Author:**

Yingjiang Ye, PhD, MD

Beijing Key Laboratory of Colorectal Cancer Diagnosis and Treatment Research

6A ward, Department of Gastroenterological Surgery, Peking University People's Hospital

No 11 Xizhimen South Street, Xicheng District

Beijing, 100044

China

Phone: 86 10 88326600

Email: [yingjiangye@pkuph.edu.cn](mailto:yingjiangye@pkuph.edu.cn)

**Related Article:**

Correction of: <https://mhealth.jmir.org/2021/3/e22647>

(*JMIR Mhealth Uhealth* 2021;9(4):e29325) doi:[10.2196/29325](https://doi.org/10.2196/29325)

In “A Novel Remote Follow-Up Tool Based on an Instant Messaging/Social Media App for the Management of Patients With Low Anterior Resection Syndrome: Pilot Prospective Self-Control Study” (*JMIR Mhealth Uhealth* 2021;9(3):e22647) the authors noted one error.

In the “Acknowledgments” section, a correction has been made to show the correct Grant Number.

The following sentence appeared in the originally published manuscript:

*This project is supported by the National Key R&D Program of China (Grant No. 2145000042).*

This sentence has been corrected to:

*This project is supported by the National Key R&D Program of China (Grant No. 2017YF0908203).*

The correction will appear in the online version of the paper on the JMIR Publications website on April 19, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 01.04.21; this is a non-peer-reviewed article; accepted 05.04.21; published 10.04.21.

*Please cite as:*

Liu F, Guo P, Su X, Cui M, Jiang J, Wang S, Yu Z, Zhou R, Ye Y

Correction: A Novel Remote Follow-Up Tool Based on an Instant Messaging/Social Media App for the Management of Patients With Low Anterior Resection Syndrome: Pilot Prospective Self-Control Study

*JMIR Mhealth Uhealth* 2021;9(4):e29325

URL: <http://mhealth.jmir.org/2021/4/e29325/>

doi: [10.2196/29325](https://doi.org/10.2196/29325)

PMID: [33872183](https://pubmed.ncbi.nlm.nih.gov/33872183/)

©Fan Liu, Peng Guo, Xiangqian Su, Ming Cui, Jianlong Jiang, Suo Wang, Zhouman Yu, Runhe Zhou, Yingjiang Ye. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 11.04.2021. 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 <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



---

Publisher:  
JMIR Publications  
130 Queens Quay East.  
Toronto, ON, M5A 3Y5  
Phone: (+1) 416-583-2040  
Email: [support@jmir.org](mailto:support@jmir.org)

---

<https://www.jmirpublications.com/>