JMIR mHealth and uHealth

Impact Factor (2022): 5.0
Volume 9 (2021), Issue 6 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

Contents

Viewpoint

Reviews

Original Papers

Corrigenda and Addenda

Viewpoint

Digital Solutions to Alleviate the Burden on Health Systems During a Public Health Care Crisis: COVID-19 as an Opportunity

Sofie H Willems[1], MD, PhD; Jyotsna Rao[1], MS; Sailee Bhambere[1], BDS, MPH; Dipu Patel[1], MPAS, PA-C; Yvonne Biggins[1], MBA, MAPP; Jessica W Guite[1,2], PhD

[1]DayToDay Health, Health Innovators Inc, Boston, MA, United States

[2]Center for Advancement in Managing Pain, School of Nursing, The University of Connecticut, Storrs, CT, United States

Corresponding Author:
Jessica W Guite, PhD
Center for Advancement in Managing Pain
School of Nursing
The University of Connecticut
231 Glenbrook Road
Storrs, CT, 06269-4026
United States
Phone: 1 215 964 5582
Email: jessica.guite@uconn.com

Abstract

The COVID-19 pandemic has generated unprecedented and sustained health management challenges worldwide. Health care systems continue to struggle to support the needs of the majority of infected individuals that are either asymptomatic or have mild symptoms. In addition, long-term effects in the form of long-lasting COVID-19 symptoms or widespread mental health issues aggravated by the pandemic pose a burden on health care systems worldwide. This viewpoint article considers aspects of digital health care solutions and how they can play an ongoing role in safely addressing gaps in the health care support available from initially and repeatedly overwhelmed providers and systems. Digital solutions can be readily designed to address this need and can be flexible enough to adapt to the evolving management requirements of various stakeholders to reduce COVID-19 infection rates, acute hospitalizations, and mortality. Multiplatform solutions provide a hybrid model of care, which can include mobile and online platforms accompanied by direct clinician input and feedback. Desirable components to be included are discussed, including symptom tracking, patient education, well-being support, and bidirectional communication between patients and clinicians. Customizable and scalable digital health platforms not only can be readily adapted to further meet the needs of employers and public health stakeholders during the ongoing pandemic, but also hold relevance for flexibly meeting broader care management needs into the future.

(JMIR Mhealth Uhealth 2021;9(6):e25021) doi:10.2196/25021

KEYWORDS

coronavirus; digital health; multiplatform; chat; symptom tracking; well-being; COVID-19; online platform; symptom; monitoring; follow-up

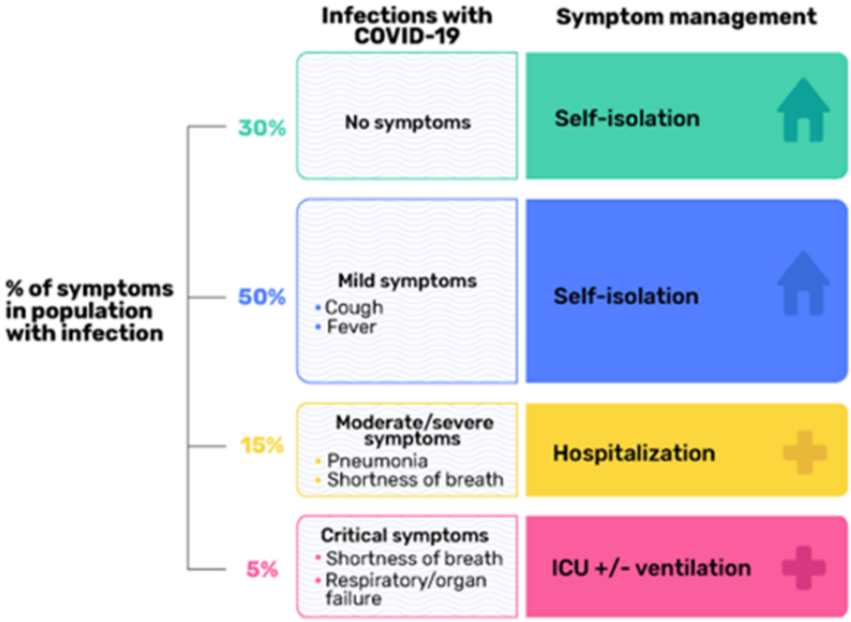
Introduction

The COVID-19 pandemic created immediate and long-term challenges for health care systems worldwide. This viewpoint examines the role that digital health care solutions play during the pandemic in safely addressing gaps in the health care support available from initially and repeatedly overwhelmed providers and systems, as well as opportunities to alleviate this burden into the future. Approximately one year into the pandemic, health care providers and systems continue to struggle with not only the care management needs of patients diagnosed with COVID-19, but also the added burden of treatment for long-lasting COVID-19 symptoms (ie, “long COVID”) [1-3] and elevated rates of mental health problems associated with the events of the past year [4,5]. Digital solutions can be readily designed to address this need and can be flexible enough to adapt to the evolving management requirements of various stakeholders. This viewpoint offers a recent historical context and perspective on how digital health technology addresses challenges presented by the COVID-19 pandemic and highlights care management opportunities for digital health solutions beyond this public health crisis.

Although the numbers are continuously changing and vary somewhat per country, initial COVID-19 research from different countries indicated that around 18%-20% of individuals who are infected have moderate to severe symptoms and require medical management [6-9]. The remaining 80% of those infected are either asymptomatic or have mild symptoms. The management of this majority subgroup—the 80% of individuals seeking medical care who had suspected or diagnosed COVID-19 with no or mild symptoms related to the disease—posed a challenge to public health management [10-13] and to stressed health care providers and systems [14-17]. Many symptomatic individuals impacted directly by COVID-19 were not hospitalized, in part due to the allocation of limited hospital resources only to those with the most severe symptoms requiring hospitalization [6,14,18]. Health care resource allocation across acute and outpatient clinical care contexts clearly presented a significant challenge in the early phase of the pandemic, which continues to strain systems of care as infection rates fluctuate.

An effective public health response to limit the spread of the virus requires those who are suspected of having COVID-19 to be tested to confirm the diagnosis and to take appropriate next steps (Figure 1). The requirement for diagnostic testing in the context of inconsistent availability of testing resources presents another layer of challenge, especially in resource-poor areas or in places where there are few diagnostic tests available. This initially resulted in large numbers of individuals who were suspected of exposure yet unable to receive timely testing, necessitating self-isolation. In addition, individuals who were positively diagnosed, but were either asymptomatic or experienced only mild to moderate symptoms, were advised to self-isolate at home with close and regular follow-up to monitor for any changes in their symptoms.

Figure 1. The burden of COVID-19 morbidity and mortality.



The infection rates among health care professionals rose in a magnitude similar to or greater than that observed in the general public [14]. Important information characterizing morbidity among essential workers at risk for COVID-19, including those health care personnel most at risk for mortality [19], emerged to support the need for continued surveillance and the development of strategies to protect essential workers and those they serve. Further underscoring this need was the imbalance between supply and demand for medical resources in many countries, which presented global questions about fair allocation of medical resources and personnel during the pandemic [6]. Factors serving to limit the pool of available health care workers included adherence to public health standards that ideally require an exposed health care worker to quarantine for 14 days. Emerging information about the risk of transmission from presymptomatic cases underscores a sustained need for vigilance, as presymptomatic individuals have an incubation period of 2-14 days, with an average of 4-5 days before COVID-19 symptoms appear and can reportedly infect 1.4-6.5 individuals during that time [20-23]. Potential provider shortages continue to be of concern during periods of high infection levels and corresponding spikes in the number of hospitalizations due to positive COVID-19 diagnoses and related diseases [16].

Such provider shortages initially necessitated reactivation of retired health care workers and onboarding of new medical personnel. For example, many US states loosened their licensing rules to give those with clinical skills the ability to participate, such as allowing out-of-state physicians to practice and requesting retired physicians to volunteer [16]. The governor of New York activated retired and student health care professionals from >52,000 volunteer health professionals, which included 2400 nurse practitioners and 2265 physicians [17]. The US Department of Veteran Affairs recruited retired federal health care workers through social media. Internationally, the Taipei Centers for Disease Control and Prevention responded

to the shortage by implementing an ongoing program to recruit and train general practitioners, retired medical professionals, and school nurses [15].

Rising rates of the disease and corresponding increases in the workload of health care personnel resulted in widespread burnout related to the escalating number of hours physicians were working without breaks [24]. In addition to treating patients in the hospital, clinicians were required to follow up on patients who were at home [12]. During this time, the mental health of health care workers was of major concern due to the effects of acute and chronic stress [19,25]. In the absence of a well-organized strategy to sustain the health care work force, scientists and clinicians called for a shift to a longer-term, more sustainable solution [26].

The uncertainty of the COVID-19 disease progression rapidly changed the landscape of the pandemic in several countries. For example, in Italy, where the burden of mortality was for a period of time relatively greater than in other countries, health care professionals perceived the response to COVID-19 as a continuously evolving process with no visible end point. In many medical settings, adapting to changes was required by the hour [27]. As the pandemic continues, new challenges have emerged, such as the management of patients experiencing "long COVID," with symptoms lasting longer than the acute infection [1,2,28]. The incidence, prevalence, and precise nature of long COVID is not entirely clear as yet [3], nor is the long-term psychological burden of COVID-19 on society as a whole. This is especially the case for patients that spent time in the intensive care unit (ICU). It is thought that more than half of ICU survivors suffer from post–intensive care syndrome (PICS), which is characterized by various symptoms including physical strength deficits, cognitive decline, and mental health disturbances observed after discharge from critical care that persist for a protracted amount of time [29,30].

According to the American Psychological Association (APA) Stress in America 2020 report and APA one-year pandemic update [4,5], there is a steadily escalating national mental and behavioral health crisis due to COVID-19. This includes negative behavioral health outcomes including unhealthy weight gain and increased drinking [4]. Nearly 1 in 5 adults (19%) say their mental health is worse than it was at this time last year and this is particularly true for the youngest generation of adults [5]. By generation, 34% of Gen Z adults report worse mental health, followed by Gen X (21%), millennials (19%), baby boomers (12%), and older adults (8%).There is a similar trend in other countries, along with notable disruptions in mental health services in most countries [31]. Overall, the psychological toll of chronic uncertainty [19] and countless other factors have created a demand for flexible solutions for disease surveillance and follow-up that could be used by workplaces and health care professionals, as well as the general public.

The Opportunity to Leverage Digital Health Solutions

The novel health care challenges emerging from the COVID-19 pandemic demanded new health care models and patient care management modalities, with digital health strategies holding great potential for delivering solutions [32]. Recent research supports how digital approaches in general can be effectively used to optimize patient care management for individuals across the life span, with attention paid to the unique developmental and psychosocial needs of adult, child, adolescent, and parent/caregiver populations, as well as those living in multigenerational family contexts [33-35]. There is emerging evidence that mobile phone apps are a feasible and acceptable way of administering health interventions for a range of chronic and acute conditions, and can possibly lead to increased patient self-management, health-related behavior change, a reduction in the use of health care resources [36,37], and improved physical and mental health outcomes [38-40]. However, most studies only showed a modest effect [38,40-43], with more and larger studies being required to determine efficacy and establish evidence for best practices. Finally, there is a small but growing economic evidence base focusing on the value of mobile health (mHealth) interventions, suggesting that they can be cost-effective, economically beneficial, and/or cost saving [44,45].

At the onset of the pandemic, there were many pre-existing popular digital apps available that could alleviate the mental health impact of the pandemic for the general public. These apps could conceivably support resilience through meditation or counseling, as well as support socially distanced care access from a health care provider through telebehavioral health. As literature with a focus on COVID-19–specific digital health solutions began to emerge, it was clear that most initial digital health solutions were oriented to surveillance and symptom tracking [46]. However, at the outset of the pandemic, there were no comprehensive digital health care pandemic control strategies identified that could address the needs of the overwhelming majority of individuals impacted by COVID-19. Solutions that could effectively support and follow up on individuals who required self-isolation due to suspected or diagnosed COVID-19 with no, mild, or moderate symptoms was identified as an important unmet need at the outset of the pandemic. As the pandemic continues into its second year, the need to support this large population of individuals who require self-isolation persists in parallel with new public health management phases that are unfolding, focusing on vaccine administration. Thus, digital health solutions continue to serve an important role in meeting the continued management needs of earlier phases of the pandemic while also simultaneously adapting to meet newly unfolding pandemic needs into the foreseeable future.

Digital Health Adaptation to Evolving Pandemic Needs

As pandemic management strategies begin to accommodate the availability of promising vaccines [47-51], the ongoing management of individuals not yet vaccinated and those who remain at risk for acquiring and transmitting SARS-CoV-2 and its new variant strains persists. This new phase of pandemic management faces many challenges, including limited and inconsistent vaccine availability as well as health inequity

factors [52] that present further barriers to achieving local, national, and global vaccination goals [49,51,53]. At the individual level, behavioral factors are critical to understanding an individual's choice to receive the vaccine or not and encouraging desired behavioral follow-through [54]. There is much to learn from behavioral scientists about critical elements to facilitate successful widespread trust and acceptance of the vaccine and the necessary follow-through for individuals to receive required vaccine doses [53,55]. These are just a few of the barriers to overcome to successfully relax COVID-19 precautions and begin to safely resume valued aspects of everyday life that were put on hold due to COVID-19.

A recently published viewpoint article by Laur and colleagues [56] provides an excellent case study of the challenges of building health services in the context of the rapidly changing COVID-19 landscape. The authors offer perspective through lessons learned in adaptive leadership, drawing upon their experience of developing a COVID-19 remote monitoring program, from a hospital-based health system perspective. They draw upon the process of making "pivots" during development, a process that is commonly used by digital health startups to manage uncertainty, and explain how these strategies hold relevance for health care leaders at any time. The viewpoint by Laur et al [56] provides an excellent consideration of the decisions made by hospital-based health service providers developing a COVID-19 remote monitoring plan and this viewpoint provides readers with additional considerations in the digital health development process.

Development of a Digital Health Solution

The key to a successful strategy for pandemic preparedness and response management is a well-planned, effectively communicated and coordinated emergency response that draws on medical mobilization. Digital health solutions are in many ways an ideal answer to meeting this need. However, they require time and resources to develop, implement, and evaluate fully. The ongoing management of COVID-19–related patient needs continues and a focus on designing a multiplatform digital health care app for health care providers to effectively and safely monitor individuals outside of the hospital setting continues to be important. Central to the process of developing a digital health solution is to ground and inform further decision making in user experience and evidence-based clinical research that speak to the importance and value of including direct input from patients and other key stakeholders to better understand their wants and needs for the digital solution. Both engagement with and adherence to use of a digital health app are critical factors to consider, particularly when there is a goal to demonstrate direct and/or indirect relationships between a patient's engagement and adherence and a desired outcome (ie, intervention effectiveness). Emerging digital health research provided insights into patient engagement [57,58] and technology-based features and strategies for promoting adherence [59,60].

Additional clinical expert input on design and development should be elicited and include the perspectives of physicians, physician assistants, and nurses. The development process also should elicit patient and caregiver input, with this feedback incorporated at various stages of development. This user experience research should be used early in the development process and include contextual, qualitative research methods to understand patients' perspective and concerns, as well as clinical, functional, and emotional needs during each phase of the patient journey. This information helps to inform and improve the patient-as-a-customer digital experience and provides a richer understanding of how patients both experience and engage with the developing app. Qualitative interviews with app users can also provide additional details about a patient's understanding of app content, its ease of use, and his/her emotional responses to the product. This information, in turn, can be used to iteratively adapt and modify the product to improve these features and to ensure both short- and longer-term engagement with the app occurs. In addition, the documentation of patient engagement and adherence information provides an important foundation for future, larger scale clinical research projects that can prospectively test the effectiveness of a digital health intervention and related research hypotheses of interest.

Clinical objectives for the design of a solution that could flexibly support social distancing guidelines intended to reduce COVID-19 infection rates, acute hospitalizations, and mortality require many additional considerations. For example, additional design considerations may include the need for it to be used flexibly by actively practicing, retired, or new medical professionals. The use of a remote, digital interface presents minimal risk of infection for health care providers, which could in turn alleviate stress and prevent further burnout. Additionally, flexibility of a digital platform to allow for changes to be readily incorporated to adapt to evolving pandemic care management needs is of great value. A design that can provide a comprehensive digital health management system that includes close safety monitoring and can deploy an escalation protocol if a patient's symptoms worsen and the patient requires acute management in a hospital setting is important. A digital platform that can be further customized to allow for maximal flexibility to adapt to the specific needs of a particular target population or treatment context is also desirable.

Consideration of Solution Components and Functionality

Platform and Staffing Considerations

The capacity to connect an individual (referred to as a patient), who is either asymptomatic with a positive diagnosis or has tested positive with mild-to-moderate symptoms, and their treating clinician is a desirable feature, particularly if symptoms escalate. To accomplish this, a design that includes both a web-based interface and an app for mobile platforms (ie, iOS and Android) is appealing. This would allow a patient to interact directly with the mobile app components on his/her smartphone and also allow for communication with a clinician. To make the monitoring process more efficient for clinicians, a combination of remote patient self-monitoring with planning for stepped levels of clinician supervision can be protocolized. Clinicians can then successfully monitor multiple patients based

on patient risk factors and symptom severity, and acuity-based staffing protocols can be flexibly designed to help to ensure that available support resources are appropriately assigned.

The Mobile App

To best serve the patient and clinician, the mobile app platform available to the patient can contain functionality for symptom management, patient education, well-being support, and communication with their health care provider, among other options. Patients can engage with educational content and well-being support provided in the mobile app, while clinicians are able to stay informed of the patient's progress through the patient's entry of vital signs and symptom tracking. This real-time data sharing would allow for timely intervention and management of patient health concerns.

Symptom Management

Symptom management can include tracking of vital signs including respiratory rate, temperature, heart rate, blood pressure, and oxygen saturation. The recorded values can be tracked over time and be made graphically visible to patients. If a vital sign is out of the appropriate physiological range, feedback is provided to the patient and can also be shared with an assigned clinician, who could then contact the patient by chat or phone. At any time, the patient could also initiate a chat conversation with the assigned clinician. A built-in daily health check can assess COVID-19 symptoms with the aim of determining the severity of the symptoms and the need for emergency medical support. Based on responses, a patient would be recommended to either continue to monitor symptoms, contact their physician, or seek emergency services. The patient could also initiate a chat functionality with an assigned clinician, who would have real-time access to the patient responses. The patients could also be provided with a temporal view of their symptoms, displayed on a symptom tracker graph, to provide perspective on progress over time to facilitate effective self-management.

Patient Education and Well-being Support

An app can be designed to include a comprehensive library of educational content for patients to access to inform themselves about COVID-19 symptoms, manage symptoms, and to support overall well-being. Content that includes evidence-based stress reduction strategies, such as mindfulness and relaxation exercises that are effective in reducing patient stress levels, can be accessible through the app [61-63]. Relaxation exercises can be presented through text, image, and video formats. These strategies should be easy to access, and are worthwhile to include as their benefits include reducing the amount of adrenaline and cortisol in the body, which are elevated in many COVID-19 patients; there are also longer term benefits when such strategies are used for dealing with stress and anxiety [64]. Even during extreme stress, exercises focused on breathing and relaxation can alleviate negative thoughts, moods, and feelings, and increase rates of recovery [65]. Additional self-management strategies including progressive muscle relation exercises, guided visualization, and a loving kindness–themed meditation can be included to support improved mental well-being [66,67]. Additional mental health screening and targeted psychological support services may also be integrated into the platform. Having this included can facilitate patients' ability to smoothly and easily engage in psychotherapeutic contact with an appropriately licensed behavioral health clinician (eg, a psychologist) when a higher level of support is indicated.

Patient-Clinician Communication

Secure, bidirectional chat communication between a patient and clinician can be initiated 24/7 through the app by the patient or through a web-based portal by the clinician. The clinician response team may be designated by a hospital or a single clinician. Once a patient initiates the chat communication functionality, the assigned clinician will receive a message and would be expected to respond within a predetermined time duration. The clinician may assess the patient's data and all data input from the patient's app-based interface in order to provide an appropriate response to the patient's query. Communication features can further support sharing images, documents, and links, as needed, and can include notifications and reminders to further enhance adherence to treatment recommendations.

Conclusion

This viewpoint provides additional perspective on how digital health solutions can be used to remotely meet the ongoing challenges that the COVID-19 pandemic presents to health care systems. Multiplatform mHealth/eHealth solutions provide an example to ground and further enrich consideration of how to best meet the evolving needs of a large, significant health care management population. Solutions designed to support the needs of both patients and clinicians safely and remotely for continued follow-up of individuals with COVID-19–related symptoms will remain necessary for the foreseeable future. Moreover, these solutions provide opportunities to flexibly support care management needs of patients, providers, and health care systems beyond this public health crisis.

Components to consider leveraging in a multiplatform digital health care solution designed to connect patients to clinicians for continued follow-up needs are discussed and can safely manage and prevent disease progression and mortality. The solution implemented should align with evidence-based recommendations such as the third domain of the Centers for Disease Control and Prevention's 2017 Pandemic Influenza Plan [13]. Domain 3 of this plan specifies medical countermeasures to increase access and use of critical countermeasures for response activities. Recently emerging literature speaks to how valuable mHealth systems and platforms that can facilitate access to mobile care providers through telehealth will continue to be for situations requiring self-isolation [46]. Moreover, there is broad-reaching potential for the utility of mHealth systems and platforms as effective solutions for future care management needs beyond the pandemic. The continually increasing widespread adoption of mobile phones worldwide presents significant opportunities for health-related apps to impact health behaviors globally, particularly in low- and middle-income countries [68].

Most solutions at the start of the pandemic were mainly intended to support contact tracing and symptom monitoring for self-use

by individuals, without any input from clinicians [46], or remote monitoring solutions created by hospitals and health service organizations adapting to rapidly address shifting institutional needs [56]. The involvement of clinicians and close monitoring of symptoms is critical for individuals with asymptomatic, mild, or moderate disease from the standpoint of the potential for the sudden emergence of severe symptoms and unexpected deterioration [69]. COVID-19–associated acute and long-term health management needs, along with continued adherence to recommended social distancing practices more generally, support the lasting value of similar digital solutions into the future.

While additional research is always needed to test the efficacy of a newly developed intervention, the urgency of the pandemic demanded a need to rapidly develop and implement digital solutions. Researchers are highlighting the need for the development of hybrid digital health solutions and models of care that combine mobile and online platforms with direct clinician input, while also addressing the needs of impacted patients across the developmental continuum [33-35,70]. Future research efforts should take steps to evaluate the efficacy of these interventions as well as the economic impact of an mHealth intervention with respect to its cost-effectiveness to add to the limited evidence base that currently exists [44,45].

The adaptability of digital platforms to flexibly accommodate various components and content can not only reduce the significant patient-care burden experienced by health care professionals, but also can be useful in other settings. For example, these digital platforms can be customized to support the needs of an employer's management of employee symptoms or the public health needs of a government continuing to combat and effectively manage the current and future phases of the COVID-19 pandemic. Ultimately, the rapid adoption of mobile technology globally [68], coupled with the adaptability of digital health platforms and their content, provides valuable opportunities for improved health care throughout the pandemic and beyond.

Acknowledgments

All authors acknowledge that they have received financial support for their work on this viewpoint manuscript from DayToDay Health, Health Innovators Inc. Many thanks to Stefan Mallette and Chien-Min Lu for their contributions to the creation of the figure and earlier versions of this manuscript, and to Christine Hsieh, PhD, for her input and steadfast support of the project.

Conflicts of Interest

None declared.

References

1. National Institute for Health and Care Excellence. Overview | COVID-19 rapid guideline: managing the long-term effects of COVID-19 | Guidance | NICE. NICE. URL: https://www.nice.org.uk/guidance/ng188 [accessed 2021-03-10]
2. The prevalence of long COVID symptoms and COVID-19 complications. Office for National Statistics, UK. URL: https://www.ons.gov.uk/news/statementsandletters/theprevalenceoflongcovidsymptomsandcovid19complications [accessed 2021-03-10]
3. Gorna R, MacDermott N, Rayner C, O'Hara M, Evans S, Agyen L, et al. Long COVID guidelines need to reflect lived experience. Lancet 2021 Feb 06;397(10273):455-457 [FREE Full text] [doi: 10.1016/S0140-6736(20)32705-7] [Medline: 33357467]
4. One year later, a new wave of pandemic health concerns. American Psychological Association. 2021. URL: https://www.apa.org/news/press/releases/stress/2021/one-year-pandemic-stress [accessed 2021-03-22]
5. Stress in America 2020: A National Mental Health Crisis. American Psychological Association. 2020 Oct. URL: https://www.apa.org/news/press/releases/stress/2020/sia-mental-health-crisis.pdf [accessed 2021-05-27]
6. Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, et al. Fair Allocation of Scarce Medical Resources in the Time of Covid-19. N Engl J Med 2020 May 21;382(21):2049-2055. [doi: 10.1056/nejmsb2005114]
7. COVID-19 Report: 19 May 2020. International Severe Acute Respiratory and Emerging Infections Consortium (ISARIC). URL: https://media.tghn.org/medialibrary/2020/05/ISARIC_Data_Platform_COVID-19_Report_19MAY20.pdf [accessed 2021-05-27]
8. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. [The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China]. Zhonghua Liu Xing Bing Xue Za Zhi 2020 Feb 10;41(2):145-151. [doi: 10.3760/cma.j.issn.0254-6450.2020.02.003] [Medline: 32064853]
9. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases 2020 Jun;20(6):669-677 [FREE Full text] [doi: 10.1016/S1473-3099(20)30243-7] [Medline: 32240634]
10. Coronavirus Disease 2019 (COVID-19). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html [accessed 2021-05-27]
11. NHS 111 online - About coronavirus (COVID-19). National Health Service. URL: https://www.nhs.uk/conditions/coronavirus-covid-19/ [accessed 2021-05-27]

12. World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected. Interim guidance. Pediatr Med Rodz 2020 Jul 14;16(1):9-26. [doi: 10.15557/pimr.2020.0003]
13. Pandemic Influenza Plan 2017 UPDATE. US Department of Health and Human Services. URL: https://www.cdc.gov/flu/pandemic-resources/pdf/pan-flu-report-2017v2.pdf [accessed 2021-05-27]
14. Adams JG, Walls RM. Supporting the Health Care Workforce During the COVID-19 Global Epidemic. JAMA 2020 Apr 21;323(15):1439-1440. [doi: 10.1001/jama.2020.3972] [Medline: 32163102]
15. Yen M, Chiu A, Schwartz J, King C, Lin Y, Chang S, et al. From SARS in 2003 to H1N1 in 2009: lessons learned from Taiwan in preparation for the next pandemic. J Hosp Infect 2014 Aug;87(4):185-193 [FREE Full text] [doi: 10.1016/j.jhin.2014.05.005] [Medline: 24996515]
16. States Get Creative To Find And Deploy More Health Workers In COVID-19 Fight. NPR. 2020 Mar 25. URL: https://www.npr.org/sections/health-shots/2020/03/25/820706226/states-get-creative-to-find-and-deploy-more-health-workers-in-covid-19-fight [accessed 2020-04-07]
17. Who Are the Retired Health Care Workers the U.S. is Calling On To Help With the Coronavirus Response? Time. URL: https://time.com/5810120/retired-health-care-workers-coronavirus/ [accessed 2020-04-07]
18. Duca LM, Xu L, Price SF, McLean CA. COVID-19 Stats: COVID-19 Incidence,* by Age Group - United States, March 1-November 14, 2020. MMWR Morb Mortal Wkly Rep 2021 Jan 01;69(5152):1664 [FREE Full text] [doi: 10.15585/mmwr.mm695152a8] [Medline: 33382674]
19. Bavel J, Baicker K, Boggio P, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. Nat Hum Behav 2020 May;4(5):460-471. [doi: 10.1038/s41562-020-0884-z] [Medline: 32355299]
20. Tang D, Comish P, Kang R. The hallmarks of COVID-19 disease. PLoS Pathog 2020 May 22;16(5):e1008536 [FREE Full text] [doi: 10.1371/journal.ppat.1008536] [Medline: 32442210]
21. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
22. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 2020 Mar 26;382(13):1199-1207 [FREE Full text] [doi: 10.1056/NEJMoa2001316] [Medline: 31995857]
23. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Ann Intern Med 2020 May 05;172(9):577-582 [FREE Full text] [doi: 10.7326/M20-0504] [Medline: 32150748]
24. 6 ways to address physician stress during COVID-19 pandemic. American Medical Association. URL: https://www.ama-assn.org/delivering-care/public-health/6-ways-address-physician-stress-during-covid-19-pandemic [accessed 2020-04-16]
25. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
26. Gold J. Surging Health Care Worker Quarantines Raise Concerns As Coronavirus Spreads. Kaiser Health News. URL: https://khn.org/news/surging-health-care-worker-quarantines-raise-concerns-as-coronavirus-spreads/ [accessed 2020-04-07]
27. COVID-19: 'A continuously evolving process that requires adapting by the hour'. Healthcare IT News. URL: https://www.healthcareitnews.com/news/europe/covid-19-continuously-evolving-process-requires-adapting-hour [accessed 2020-04-07]
28. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. BMJ 2020 Aug 11;370:m3026. [doi: 10.1136/bmj.m3026] [Medline: 32784198]
29. Elliott D, Davidson JE, Harvey MA, Bemis-Dougherty A, Hopkins RO, Iwashyna TJ, et al. Exploring the Scope of Post–Intensive Care Syndrome Therapy and Care. Critical Care Medicine 2014;42(12):2518-2526. [doi: 10.1097/ccm.0000000000000525]
30. Jaffri A, Jaffri UA. Post-Intensive care syndrome and COVID-19: crisis after a crisis? Heart Lung 2020 Nov;49(6):883-884 [FREE Full text] [doi: 10.1016/j.hrtlng.2020.06.006] [Medline: 32690219]
31. COVID-19 disrupting mental health services in most countries, WHO survey. World Health Organization. URL: https://www.who.int/news/item/05-10-2020-covid-19-disrupting-mental-health-services-in-most-countries-who-survey [accessed 2021-03-10]
32. Fagherazzi G, Goetzinger 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] [Medline: 32501804]
33. 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] [Medline: 32540841]
34. Serlachius A, Badawy S, Thabrew H. Psychosocial Challenges and Opportunities for Youth With Chronic Health Conditions During the COVID-19 Pandemic. JMIR Pediatr Parent 2020 Oct 12;3(2):e23057 [FREE Full text] [doi: 10.2196/23057] [Medline: 33001834]

35. Radovic A, Badawy SM. Technology Use for Adolescent Health and Wellness. Pediatrics 2020 May;145(Suppl 2):S186-S194 [FREE Full text] [doi: 10.1542/peds.2019-2056G] [Medline: 32358210]
36. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014 Apr 16;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
37. Merchant R, Szefler SJ, Bender BG, Tuffli M, Barrett MA, Gondalia R, et al. Impact of a digital health intervention on asthma resource utilization. World Allergy Organ J 2018;11(1):28 [FREE Full text] [doi: 10.1186/s40413-018-0209-0] [Medline: 30524644]
38. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR mHealth uHealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
39. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review. JMIR mHealth uHealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
40. 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] [Medline: 28428157]
41. 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] [Medline: 31692183]
42. 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] [Medline: 30026178]
43. 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] [Medline: 26831740]
44. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
45. Badawy SM, Kuhns LM. Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review. JMIR mHealth uHealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
46. John Leon Singh H, Couch D, Yap K. Mobile Health Apps That Help With COVID-19 Management: Scoping Review. JMIR Nurs 2020 Aug 6;3(1):e20596 [FREE Full text] [doi: 10.2196/20596] [Medline: 32897271]
47. COVID-19 and Your Health: Different Vaccines. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines.html [accessed 2021-01-27]
48. Draft landscape and tracker of COVID-19 candidate vaccines. World Health Organization; WHO Team: R&D Blue Print. URL: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines [accessed 2021-01-27]
49. Chung JY, Thone MN, Kwon YJ. COVID-19 vaccines: The status and perspectives in delivery points of view. Adv Drug Deliv Rev 2021 Mar;170:1-25 [FREE Full text] [doi: 10.1016/j.addr.2020.12.011] [Medline: 33359141]
50. Jeyanathan M, Afkhami S, Smaill F, Miller MS, Lichty BD, Xing Z. Immunological considerations for COVID-19 vaccine strategies. Nat Rev Immunol 2020 Oct 04;20(10):615-632 [FREE Full text] [doi: 10.1038/s41577-020-00434-6] [Medline: 32887954]
51. Kaur SP, Gupta V. COVID-19 Vaccine: A comprehensive status report. Virus Res 2020 Oct 15;288:198114 [FREE Full text] [doi: 10.1016/j.virusres.2020.198114] [Medline: 32800805]
52. National Academies of Sciences, Engineering, and Medicine. Framework for Equitable Allocation of COVID-19 Vaccine. Washington, DC: The National Academies Press; 2020. [doi: 10.17226/25917]
53. National Governors Association, Duke-Margolis Center for Health Policy, COVID Collaborative. Supporting an Equitable Distribution of COVID-19 Vaccines: Key Themes, Strategies, and Challenges Across State and Territorial COVID-19 Vaccination Plans. URL: https://healthpolicy.duke.edu/sites/default/files/2020-12/Supporting%20an%20Equitable%20Distribution%20of%20COVID-19%20Vaccines%20FINAL.pdf [accessed 2021-05-27]
54. Kritz F. 'Trusted Messengers, Trusted Messages': How To Overcome Vaccine Hesitancy. NPR. URL: https://www.npr.org/sections/health-shots/2020/12/24/948776228/trusted-messengers-trusted-messages-how-to-overcome-vaccine-hesitancy [accessed 2021-01-27]
55. Dooling K, Marin M, Wallace M, McClung N, Chamberland M, Lee GM, et al. The Advisory Committee on Immunization Practices' Updated Interim Recommendation for Allocation of COVID-19 Vaccine - United States, December 2020. MMWR Morb Mortal Wkly Rep 2021 Jan 01;69(5152):1657-1660 [FREE Full text] [doi: 10.15585/mmwr.mm695152e2] [Medline: 33382671]

56. Laur CV, Agarwal P, Mukerji G, Goulbourne E, Baranek H, Pus L, et al. Building Health Services in a Rapidly Changing Landscape: Lessons in Adaptive Leadership and Pivots in a COVID-19 Remote Monitoring Program. J Med Internet Res 2021 Jan 13;23(1):e25507 [FREE Full text] [doi: 10.2196/25507] [Medline: 33417588]
57. 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] [Medline: 27966189]
58. 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] [Medline: 28241759]
59. Badawy SM, Thompson AA, Kuhns LM. Medication Adherence and Technology-Based Interventions for Adolescents With Chronic Health Conditions: A Few Key Considerations. JMIR mHealth uHealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
60. 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] [Medline: 26844685]
61. van Dixhoorn J, White A. Relaxation therapy for rehabilitation and prevention in ischaemic heart disease: a systematic review and meta-analysis. Eur J Cardiovasc Prev Rehabil 2005 Jun;12(3):193-202. [doi: 10.1097/00149831-200506000-00002] [Medline: 15942415]
62. Neves A, Alves A, Ribeiro F, Gomes J, Oliveira J. The effect of cardiac rehabilitation with relaxation therapy on psychological, hemodynamic, and hospital admission outcome variables. J Cardiopulm Rehabil Prev 2009;29(5):304-309. [doi: 10.1097/HCR.0b013e3181b4ca27] [Medline: 19935143]
63. Klainin-Yobas P, Oo WN, Suzanne Yew PY, Lau Y. Effects of relaxation interventions on depression and anxiety among older adults: a systematic review. Aging Ment Health 2015;19(12):1043-1055. [doi: 10.1080/13607863.2014.997191] [Medline: 25574576]
64. Jerath R, Crawford MW, Barnes VA, Harden K. Self-regulation of breathing as a primary treatment for anxiety. Appl Psychophysiol Biofeedback 2015 Jun;40(2):107-115. [doi: 10.1007/s10484-015-9279-8] [Medline: 25869930]
65. Broadbent E, Kahokehr A, Booth RJ, Thomas J, Windsor JA, Buchanan CM, et al. A brief relaxation intervention reduces stress and improves surgical wound healing response: a randomised trial. Brain Behav Immun 2012 Feb;26(2):212-217. [doi: 10.1016/j.bbi.2011.06.014] [Medline: 21741471]
66. Antall GF, Kresevic D. The use of guided imagery to manage pain in an elderly orthopaedic population. Orthop Nurs 2004;23(5):335-340. [doi: 10.1097/00006416-200409000-00012] [Medline: 15554471]
67. Huddleston P. Prepare for Surgery, Heal Faster: A Guide of Mind-body Techniques. Cambridge, MA: Angel River Press; 1996.
68. 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]
69. Jarrett M, Schultz S, Lyall J, Wang J, Stier L, De Geronimo M, et al. Clinical Mortality in a Large COVID-19 Cohort: Observational Study. J Med Internet Res 2020 Sep 25;22(9):e23565 [FREE Full text] [doi: 10.2196/23565] [Medline: 32930099]
70. Balcombe L, De Leo D. An Integrated Blueprint for Digital Mental Health Services Amidst COVID-19. JMIR Ment Health 2020 Jul 22;7(7):e21718 [FREE Full text] [doi: 10.2196/21718] [Medline: 32668402]

Abbreviations

APA: American Psychological Association
ICU: intensive care unit
mHealth: mobile health
PICS: post–intensive care syndrome

Edited by L Buis; submitted 15.10.20; peer-reviewed by S Badawy, MD, MS, G Boyle; comments to author 02.12.20; revised version received 01.02.21; accepted 18.05.21; published 11.06.21.

Please cite as:
Willems SH, Rao J, Bhambere S, Patel D, Biggins Y, Guite JW
Digital Solutions to Alleviate the Burden on Health Systems During a Public Health Care Crisis: COVID-19 as an Opportunity
JMIR Mhealth Uhealth 2021;9(6):e25021
URL: https://mhealth.jmir.org/2021/6/e25021
doi:10.2196/25021
PMID:34033575

Review

Factors That Influence the Use of Electronic Diaries in Health Care: Scoping Review

Naomi E M Daniëls[1,2], MSc; Laura M J Hochstenbach[3], PhD; Catherine van Zelst[1], PhD; Marloes A van Bokhoven[2], MD, PhD; Philippe A E G Delespaul[1,4], PhD; Anna J H M Beurskens[2], PhD

[1]Department of Psychiatry and Neuropsychology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands
[2]Department of Family Medicine, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands
[3]Research Centre for Remote Health Care, Faculty of Health Care, Zuyd University of Applied Sciences, Heerlen, Netherlands
[4]Mondriaan Mental Health Trust, Heerlen/Maastricht, Netherlands

Corresponding Author:
Naomi E M Daniëls, MSc
Department of Psychiatry and Neuropsychology
Faculty of Health, Medicine and Life Sciences
Maastricht University
Vijverdalseweg 1
Maastricht, 6226NB
Netherlands
Phone: 31 43 3883820
Email: naomi.daniels@maastrichtuniversity.nl

Abstract

Background: A large number of people suffer from psychosocial or physical problems. Adequate strategies to alleviate needs are scarce or lacking. Symptom variation can offer insights into personal profiles of coping and resilience (detailed functional analyses). Hence, diaries are used to report mood and behavior occurring in daily life. To reduce inaccuracies, biases, and noncompliance with paper diaries, a shift to electronic diaries has occurred. Although these diaries are increasingly used in health care, information is lacking about what determines their use.

Objective: The aim of this study was to map the existing empirical knowledge and gaps concerning factors that influence the use of electronic diaries, defined as repeated recording of psychosocial or physical data lasting at least one week using a smartphone or a computer, in health care.

Methods: A scoping review of the literature published between January 2000 and December 2018 was conducted using queries in PubMed and PsycInfo databases. English or Dutch publications based on empirical data about factors that influence the use of electronic diaries for psychosocial or physical purposes in health care were included. Both databases were screened, and findings were summarized using a directed content analysis organized by the Consolidated Framework for Implementation Research (CFIR).

Results: Out of 3170 articles, 22 studies were selected for qualitative synthesis. Eleven themes were determined in the CFIR categories of intervention, user characteristics, and process. No information was found for the CFIR categories inner (eg, organizational resources, innovation climate) and outer (eg, external policies and incentives, pressure from competitors) settings. Reminders, attractive designs, tailored and clear data visualizations (intervention), smartphone experience, and intrinsic motivation to change behavior (user characteristics) could influence the use of electronic diaries. During the implementation process, attention should be paid to both theoretical and practical training.

Conclusions: Design aspects, user characteristics, and training and instructions determine the use of electronic diaries in health care. It is remarkable that there were no empirical data about factors related to embedding electronic diaries in daily clinical practice. More research is needed to better understand influencing factors for optimal electronic diary use.

(JMIR Mhealth Uhealth 2021;9(6):e19536) doi:10.2196/19536

KEYWORDS

compliance; delivery of health care; diary; ecological momentary assessment; intention; motivation; scoping review

Introduction

Health care professionals are insufficiently aware of symptom variability and contextual fluctuations; therefore, their interventions are based on incomplete information [1-5]. Patients are asked to recall their mood, thoughts, behavior, and experiences over the past weeks or even months. Recalling information from memory, though, is known to be incomplete and inaccurate [6,7]. To minimize inaccuracies and biases, prospective diaries are used to collect patients' mood, thoughts, behavior, and experiences in the relevant context close to the time of occurrence [8]. Because these health-related strategies often require management of vulnerabilities, long-term patient engagement is important. However, patients experience that it is difficult to be engaged in the use of diaries for long periods of time. Compliance is often poor, and adequate reports on contextual variation are lacking [8]. Paper diaries are remarkably completed in the parking lot before meeting the clinician [9]. In one-third of the days, paper diaries contain entries while the log booklets were not opened [8,10].

To overcome noncompliance with paper diaries, researchers and clinicians have shifted from paper to electronic diaries. Both paper and electronic diaries can be used in research to observe individuals in their context, gather data about sensitive topics, or to actively engage individuals in monitoring and reflecting on behaviors, their underlying mechanisms, and processes. Furthermore, these diaries can be implemented in intervention studies, clinical trials, and routine care [11,12]. Electronic diaries are, however, more reliable and logistically easier to implement [13,14]. They allow individuals to monitor in daily life with little retrospection and reduced obtrusiveness. Electronic diaries are signal-contingent and often record response-time information, which improves reliability [15-18]. Nonetheless, electronic diaries also have disadvantages. Development and maintenance are costly [12]. Technical problems occur, and not all patients are acquainted with smartphones and require instructions and coaching [15]. Furthermore, research on compliance is ambiguous. For instance, the percentage of completed diary entries with electronic diaries ranges from less than 50% to 99% [18-20]. High participant motivation is related to accurate data collection and less faked compliance [13].

Previous research states that various factors are related to the use of electronic diaries, such as the design (ie, ease of use, entertainment value), the social context (ie, satisfaction and connection with others), and the user's characteristics (ie, education and self-efficacy) [21-23]. However, no complete overview is available concerning empirical data about the factors related to the use of these tools. Therefore, the main aim of this paper was to map the existing empirical knowledge about the factors that influence the use of electronic diaries in health care. Electronic diaries in health care were defined as repeated individual psychosocial or physical data collection using measurement tools on a smartphone (applications) or on a computer (website), including among others, experience sampling, ambulatory assessment, and ecological momentary assessment. In addition, use was defined as the repeated recording of information in electronic diaries by patients or healthy individuals for at least one week, including adherence, compliance, and engagement. The cut-off point was determined based on the expected recall bias and necessary data for comprehensive functional diagnostics.

Methods

In order to map existing knowledge concerning the topic of interest and to identify any gaps, this scoping review was based on the methodological framework proposed by Arksey and O'Malley [24]. This framework includes 5 specific steps: identify the research question, identify relevant studies, select relevant studies, chart the data, and summarize and report the results. The selection of relevant studies was not based on methodological quality, but on relevance.

Identify the Research Question

The research question of this scoping review was based on prior research and the expertise of the research team. It is summarized as: "What is the current empirical knowledge regarding factors that influence the use of electronic diaries in health care?"

Identify and Select Relevant Studies

A structured literature search was conducted using the PubMed and PsycInfo databases to search for articles published between 2000 and 2018. The search was limited to human adults and articles published in Dutch or English. Both free-text search terms and MESH headings were used. The search strategy included 2 different concepts: "continued use" and "electronic diaries." The search string used is depicted in Textbox 1. In addition to the database search, reference lists of relevant studies were screened manually for further relevant papers. This is a valuable step (snowball method) to identify articles that have been missed in the database search because electronic databases may be incomplete and they can vary in coverage, indexing, and depth of information [24]. Moreover, 2 experts in the field were contacted to identify key authors or key publications on the topic of interest.

Textbox 1. Search string.

1. Use: "compliance (MeSH) OR intention (MeSH) OR motivation (MeSH) OR 'continued usage' OR use OR continuance OR adherence OR engagement";

AND

2. Electronic diaries: "momentary (MeSH: ecological momentary assessment) OR 'real time data' (MeSH) OR e-diaries OR electronic diar* OR structured diar* OR computer diar* OR 'experience sampling' OR ambulatory assessment OR electronic assessment* OR electronic interview* OR self-monitoring"

Limits:

- Publication date: 2000-2018
- Humans: adult
- Language: English, Dutch

Two researchers (NEMD, LMJH) reviewed the retrieved studies using a 3-step screening process: titles, abstracts, and full articles. The screening process of a scoping review is not linear but rather iterative, which required the researchers to engage with each step in a reflexive way and repeat steps to ensure that the literature was covered in an extensive way. If the relevance of a study was unclear from the title, the abstract was ordered, and if the relevance of a study was unclear from the abstract, the full article was ordered. As a check on the 3-step screening process, we read the full texts of a random sample of 50 titles and 50 abstracts. In only 4 articles, we found information in the results or the discussion related to our scope. Relevant studies with the following criteria were included: (1) using electronic diaries for psychosocial or physical data, (2) describing factors that influence the use of electronic diaries, and (3) a focus on health care. No methodological criteria were applied, and articles based on empirical data were included. Studies were excluded when the definitions of electronic diaries or use in the article did not match with the ones used in this manuscript (ie, the data collection method: single moment data collection or passive self-monitoring using sensors, activity trackers, or biomarkers). Studies that used a combination of active and passive monitoring were not excluded. Moreover, studies were excluded when the article did not include factors that influence the use of electronic diaries as the outcome (ie, the study aim: experiences with disease management, epidemiology, health technology assessment, prediction models, outcome and effect studies, and the study design [reviews, secondary analysis, protocols]). Studies in which disease management were based on or complemented with self-reporting and studies about technology acceptance were not excluded. Furthermore, we excluded studies with a target population other than adults. For children and adolescents, we expect that different factors influence the use of electronic diaries specifically and interventions in general as parents, for instance, need to give their permission. At each step, the articles were categorized as relevant, irrelevant, and dubious according to the aforementioned exclusion criteria. Differences were discussed until consensus was reached. When no consensus was reached or questions remained, a third researcher (CvZ) was consulted.

Chart the Data

The data were charted using Excel spreadsheets and included study details (author, title, database, journal, year of publication, study location [published and conducted], study population and sample size, study aims, design, and setting), intervention characteristics (aim, content, and duration of the electronic diary), and key findings (factors that influence the use). These factors were organized according to the Consolidated Framework for Implementation Research (CFIR) [25]. This framework consists of 5 categories (ie, intervention characteristics, outer setting, inner setting, individual characteristics, and process) related to sustainable implementation. The intervention characteristics category includes, among others, the complexity of the electronic diary or the ability to test the electronic diary on a small scale. The outer setting category is comprised of the economic, political, and social context of the organization. The inner setting category includes, among others, the internal architecture of the organization and the innovation climate. The individual characteristics category is comprised of, among others, the individual's knowledge, beliefs, and self-efficacy regarding the intervention or the implementation process. The process category includes activities (planning, engaging, executing, reflecting, and evaluating) related to the implementation process.

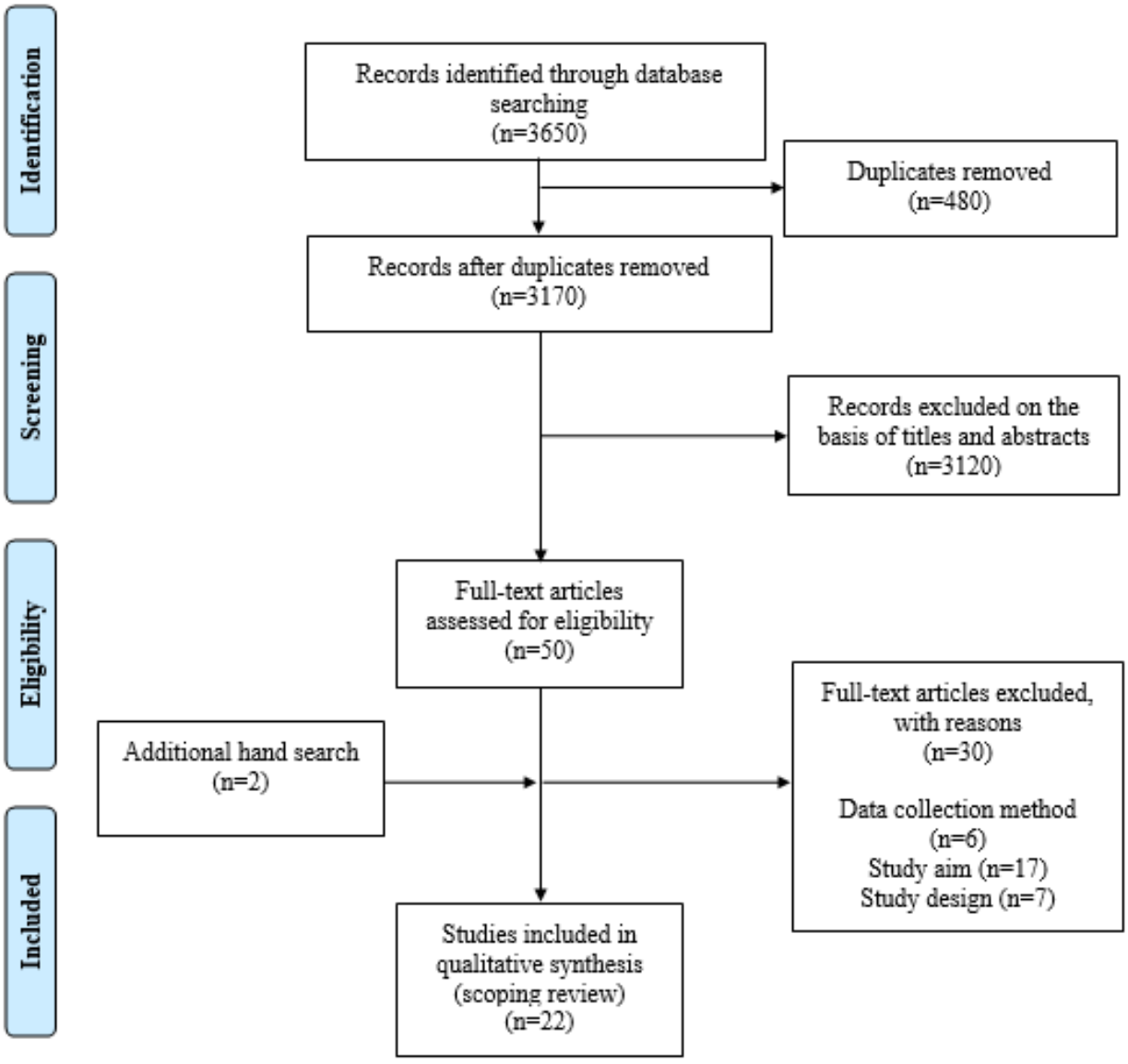
Summarize and Report the Results

Content analysis was done independently by 2 reviewers (NEMD, LMJH) based on the 5 categories of the CFIR [25]: (1) intervention, (2) outer setting, (3) inner setting, (4) individual characteristics, and (5) process. Directed content analysis, using inductive reasoning, was used to validate or conceptually extend the framework [26]. The themes were based on our previous work [27] and emerged from the data. After coding, the researchers compared their codes until consensus was reached. They identified key themes into which the results could be divided.

Results

The database search resulted in 3650 hits (Figure 1). After removing duplicates and reviewing 3170 titles, 273 abstracts were screened, of which 50 full texts were evaluated. In total, 20 articles were included based on the predefined eligibility criteria. Two articles were included from the additional hand search, which resulted in 22 articles in total for qualitative synthesis. The publication patterns are summarized in Multimedia Appendix 1.

Figure 1. Scoping review flow diagram.



Electronic Diary and Study Characteristics

More detailed information about the content of the 22 selected studies with empirical data on factors that influence the use of electronic diaries in health care can be found in Table 1. Electronic diaries were used either to monitor one's own behavior in order to get insight into underlying patterns or mechanisms (monitoring: 12/22, 55%) or to actively achieve change (intervention: 10/22, 45%). They mainly focused on measuring lifestyle behaviors (14/22, 64%) and constructs such as pain or mood. Participants completed these electronic diaries via palmtop (3/22, 14%), smartphone (14/22, 64%), or (tablet) computer (5/22, 22%). The assessment frequency ranged from 12 times a day, an example of the experience sampling method or ecological momentary assessment (EMA), to weekly, and the duration of the data collection varied from 2 weeks to 2.5 years.

Table 1. Electronic diary (e-diary) and study characteristics.

First author, year, country	e-Diary characteristics			Study characteristics		
	Purpose of use[a] (device)	Constructs measured	Frequency of use and duration	Study aims	Design and data collection	Sample: target population, number of participants, sex, age (years)
Aaron, 2004 [28], US	Intervention: cognitive behavioral therapy-based pain management training (palmtop)	Pain intensity, pain-related activity interference, jaw use limitations, mood, perceived stress	3 times a day for 8 weeks	Self-reported reasons for missing electronic diary interviews (EMA[b])	Quantitative: secondary analysis of existing RCT[c] data (CBT[d]-based pain management training or self-care manual condition)	Patients with TMD[e] (n=62), 16% male (n=10), mean age 38.6 (SD 11.6)
Litcher-Kelly, 2007 [29], US	Monitoring: self-monitoring diaries (palmtop)	Mood, stress, pain, medication use	12 times a day for 3 weeks	Feasibility of an electronic diary	Quantitative: intervention study with continuous log data	Patients with inflammatory bowel disease (n=16), 25% male (n=4), mean age 46.0 (SD 13.6)
Welch, 2007 [30], US	Monitoring: self-monitoring diaries (palmtop)	Food and fluid intake	3 times a day for 12 weeks	Feasibility of electronic self-monitoring diaries	Quantitative: pilot study with surveys	Patients on hemodialysis (n=3), 67% male (n=2), mean age 54
Stevens, 2008 [31], US	Intervention: IT[f] weight loss program (computer)	Weight, food records, exercise minutes	Weekly for a 2.5-year follow-up	First year utilization and development process of an IT weight loss program	Quantitative: RCT with 3 groups (no-further treatment, control condition, or active maintenance weight loss intervention)	Adults with a BMI of 25-45 kg/m^2 who were taking medication for hypertension or hyperlipidemia (n=348), 37% male (n=128), mean age 56
Webber, 2010 [32], US	Monitoring: internet behavioral weight loss program (computer)	Daily caloric intake, daily exercise, weight	At least weekly for 16 weeks	Motivation and adherence to self-monitoring and weight loss	Quantitative: secondary analysis of existing RCT data (did or did not achieve 5% weight loss)	Adult women with a BMI of 25-40 kg/m^2 (n=66), mean age 50.1 (SD 9.9)
Ahtinen, 2013 [33], Finland	Intervention: Oiva, a mobile mental wellness training application (smartphone)	Reflections and notes on exercises	Daily for a month	Use, acceptance, and usefulness of Oiva	Mixed methods: feasibility study with surveys, app log data and interviews	Individuals interested in stress management (n=15), 40% male (n=6), working age
Ben-Zeev, 2013 [34], US	Intervention: FOCUS, a mobile illness self-management system (smartphone)	Medication adherence, mood regulation, sleep, social functioning, coping with persistent auditory hallucinations	Daily	Development of FOCUS	Mixed methods: usability study with surveys and think-aloud procedure	Patients with schizophrenia or schizoaffective disorder (n=12), 67% male (n=8), mean age 45

First author, year, country	e-Diary characteristics			Study characteristics		
	Purpose of use[a] (device)	Constructs measured	Frequency of use and duration	Study aims	Design and data collection	Sample: target population, number of participants, sex, age (years)
Ma, 2013 [35], US	Intervention: eHealth weight loss intervention (computer)	Weight, physical activity	12 weeks, no app use criteria	Acceptance and use of an eHealth weight management intervention	Quantitative: secondary analysis of existing RCT data (coach-led or self-directed group)	Overweight or obese adults with prediabetes and/or metabolic syndrome (n=133), 53% male (n=70), mean age 53.5 (SD 10.5)
Tatara, 2013 [36], Norway	Monitoring: Few Touch, a mobile self-management application (smartphone)	Nutritional habits	1 year, no app use criteria	Factors associated with use of Few Touch, a mobile self-management application	Mixed methods: longitudinal intervention trial with surveys, interviews, and focus groups	Individuals with type 2 diabetes mellitus (n=12), 33% male (n=4), mean age 55.1 (SD 9.6)
Tang, 2015 [37], UK	Monitoring: publicly available free applications MyFitness Pal, Livestrong, Calorie Count, SparkPeople (smartphone)	Not specified	3 weeks, no app use criteria	Understanding of users' experiences with weight loss or weight control apps	Qualitative: semistructured interviews	Young adults having experience with or interest in using an eHealth weight loss maintenance app (n=19), 54% male (n=10), age range 19-33
Triantafyllidis, 2015 [38], UK	Monitoring: SUPPORT-HF, a remote health monitoring and nonpharmacological, self-monitoring system (tablet computer)	Physiological measurements (blood pressure, weight, oxygen saturation), heart failure symptoms, quality of life	5 days a week for 1 year	Development of SUPPORT-HF	Mixed methods: iterative refinement approach informed by action research	Patients with heart failure (n=26), 65% male (n=17), mean age 72 (SD 15)
Anderson, 2016 [39], Australia	Monitoring: applications about chronic conditions (sleep disorders, migraine, menstrual irregularities, chronic depression, arthritis and Behçet's disease; smartphone)	Ranging from symptom monitoring or management apps to fitness apps	Ranging from several weeks to 2 years	Consumers' experiences with mobile health apps	Qualitative: individual semistructured interviews	Healthy individuals reporting the recent use of any commercially available health/fitness app with capacity for self-monitoring and data input (n=22), 32% male (n=7), age range 18-55

First author, year, country	e-Diary characteristics			Study characteristics		
	Purpose of use[a] (device)	Constructs measured	Frequency of use and duration	Study aims	Design and data collection	Sample: target population, number of participants, sex, age (years)
Batink, 2016 [40], The Netherlands	Intervention: ACT-DL, a mobile acceptance and commitment therapy in daily life training (smartphone)	Sleep quality, appraisal of the day, affect (positive and negative feelings), cognition, context (activity, company and whereabouts)	10 times a day for 3 days each week, for 4 weeks	Feasibility, acceptability, and effectiveness of ACT-DL (EMI[g])	Mixed methods: intervention study with 2 groups (experimental intervention or outpatient treatment)	Patients with a mental health disorder such as anxiety, mood, somatoform, or substance disorders: experimental intervention (n=49), 35% male (n=17), mean age 45.7 (SD 10.0); healthy individuals (n=112), 55% male (n=62), mean age 47.5 (SD 12.4)
Jiang, 2016 [41], US	Monitoring: Pocket Personal Assistant for Training Health (Pocket PATH), a health self-monitoring application (smartphone)	Spirometry, temperature, blood pressure, pulse, symptoms, weight	12 months posttransplantation, no app use criteria	Acceptance and use of Pocket PATH	Quantitative: cross-sectional correlational design with secondary analysis of existing RCT data	Lung transplantation recipients transferred to the acute cardiothoracic unit (n=96), 51% male (n=49), mean age 57 (SD 14)
Naughton, 2016 [42], UK	Intervention: Q-sense, a smoking cessation mobile phone application (smartphone)	Smoking behavior, psychological context, situational context	1 month before until 2 weeks after a preset quit date	Feasibility of Q-sense (EMI)	Mixed methods: an explanatory sequential mixed methods design with app log data and semistructured interviews	Adult smokers willing to set a quit date in the period between 1 week and 1 month after inclusion (n=15), 53% male (n=8), age range 18-45
Timmerman, 2016 [43], The Netherlands	Monitoring: telehealth care application with a symptom monitoring module and web-based exercise module (smartphone and computer)	Pain, fatigue, dyspnea	3 days a week during 2 weeks presurgery, the first month postsurgery, and 2 weeks prior to the doctor consultation at 3 and 6 months postsurgery	Development and usability of a multimodal ICT[h]-supported rehabilitation program for lung cancer	Qualitative: user-centered design with interviews and focus groups	Patients with NSCLC[i] (n=10), 30% male (n=3), mean age 62 (SD 11)
Burke, 2017 [44], US	Intervention: standard behavioral intervention for weight (smartphone)	Not specified	5 times a day for 12 months	Lessons learned from development and implementation of an EMA study, focusing on the methods and logistics of conducting an EMA study and including strategies to ensure adequate adherence to EMA prompts	Qualitative: single-group, observational design	Former participants of laboratory weight loss studies (n=133), 9% male (n=12), mean age 51.09 (SD 10.10)

First author, year, country	e-Diary characteristics			Study characteristics		
	Purpose of use[a] (device)	Constructs measured	Frequency of use and duration	Study aims	Design and data collection	Sample: target population, number of participants, sex, age (years)
Crane, 2017 [45], UK	Monitoring: DrinkLess, an application (smartphone)	Consequences of alcohol consumption, mood, productivity, clarity, sleep quality	Daily, at least 2 weeks, no app use criteria	Usability of DrinkLess	Qualitative: usability studies with think-aloud procedure and semistructured interviews	Healthy individuals (n=12 for both studies), 50% male (n=6), mean age 42 (first study) and 40 (second study)
Freyne, 2017 [46], Australia	Intervention: PMRP[j], a behavioral-based mobile weight management program and application (smartphone)	Meal diary for previous day, current weight, dietary intake, update food diary	3 times a day for an intervention period of 12 weeks, followed by another 12-week period	Role of push notifications in persuading users to engage with self-monitoring tasks	Quantitative: intervention study with app log data	Overweight adults (BMI >25 kg/m^2; n=75), 27% male (n=20), mean age 48.6
Kreyenbuhl, 2018 [47], US	Intervention: MedActive, an application (smartphone)	Medication adherence, positive psychotic symptoms, medication side effects	Daily for 2 weeks	Acceptability and feasibility of MedActive (EMA)	Quantitative: user-centered design with surveys	Patients with schizophrenia spectrum disorder taking ≥1 oral antipsychotic medications (n=7), 100% male (n=7), mean age 47.6 (SD 10.4)
Liu, 2018, US [48]	Monitoring: LoseIt, a physical activity and diet tracking application (smartphone)	Food intake	At least 3 days a week for 2 weeks	Effectiveness of LoseIt	Quantitative: randomized trial with 2 groups (goal setting reminders or generic reminders) with pre- and posttests	College students (n=50), 38% male (n=19), mean age 21 (SD 1.8)
Tomko, 2018 [49], US	Monitoring: REDCap, ambulatory assessment software (computer)	Smoking, substance use, medication adherence	3 times daily for 8 weeks	Feasibility of ambulatory assessment (here applied in smoking cessation) for research purposes (EMA)	Quantitative: feasibility study within a double-blind RCT with 2 groups (N-acetylcysteine or placebo)	Adult smokers (n=36), 50% male (n=18), mean age 41.1 (SD 12.7)

[a]The purpose of use category is based on the authors' interpretation of the described goal of the electronic diary.

[b]EMA: ecological momentary assessment.

[c]RCT: randomized controlled trial.

[d]CBT: cognitive behavioral therapy.

[e]TMD: temporomandibular disorder.

[f]IT: information technology.

[g]EMI: ecological momentary intervention.

[h]ICT: information communication technology.

[i]NSCLC: non-small cell lung cancer.

[j]PMRP: partial meal replacement program.

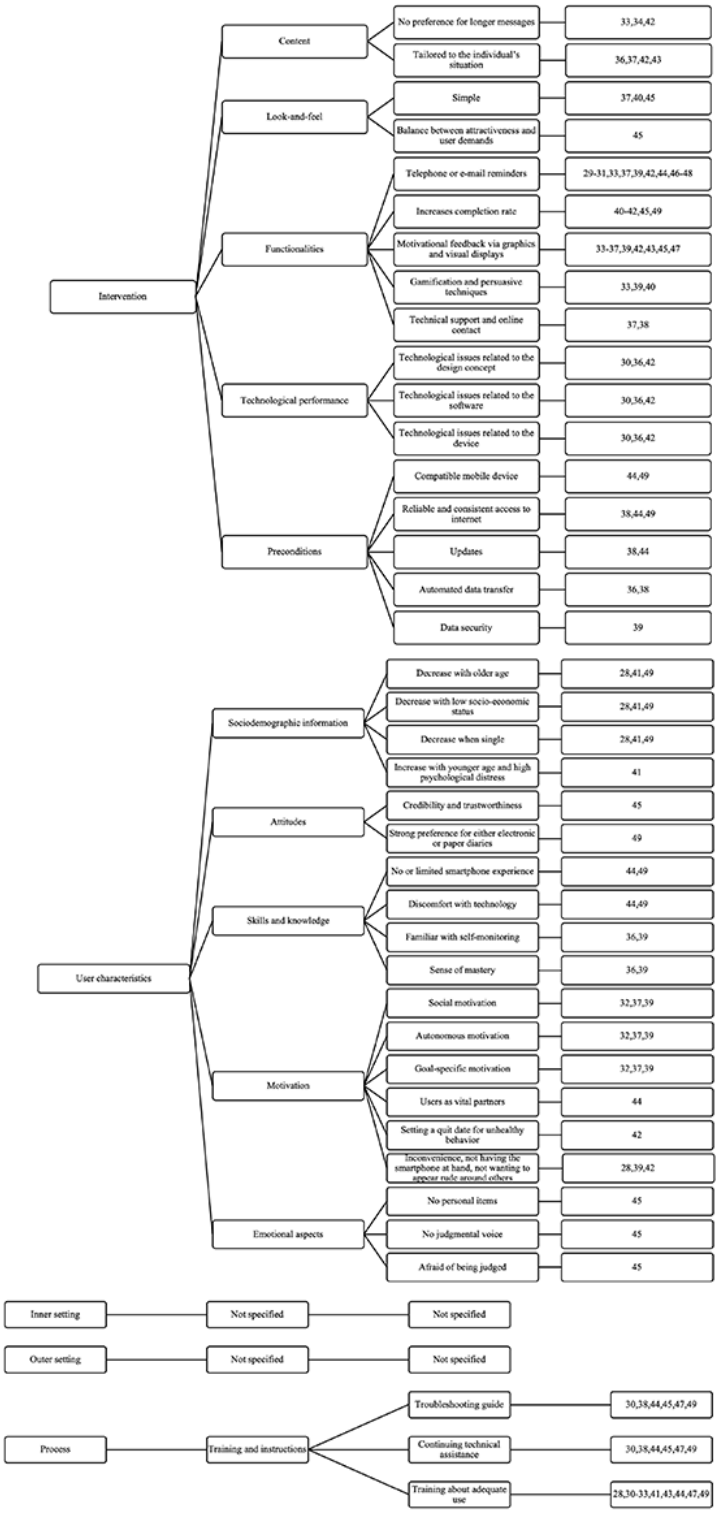
The factors that influence the use of electronic diaries in health care were not the primary aim in all included studies. These factors were mentioned as part of a larger study, such as a randomized controlled trial or an intervention study. Studies focused on usability in half of the articles (10/22, 45%), followed by feasibility and effectiveness (7/22, 32%) and development (5/22, 23%). The design of these studies was quantitative (11/22, 50%), mixed (6/22, 27%), or qualitative (5/22, 23%). The number of participants ranged from 3 to 348, with a mean age of 49 years. Of these, 37.0% (493/1341) were male. The majority of the studies included patients with physical symptoms (12/22, 55%), whereas healthy individuals (7/22,

32%) and patients with mental health symptoms (3/22, 13%) were less often described.

Factors That Influence the Use of Electronic Diaries

The CFIR [25] was used to perform the qualitative thematic analysis of the factors that influence the use of electronic diaries in health care. The results of this qualitative thematic analysis were organized along 3 CFIR categories: intervention [29-31,33-49], user characteristics [28,32,36,37,39,41,42,44,45,49], and process [30-33,38,41,43-45,47,49]. No results were found for the 2 other CFIR categories: inner setting and outer setting. Figure 2 gives an overview of these categories, themes, and subthemes.

Figure 2. Visual representation of the factors that influence the use of electronic diaries in health care.



Intervention

The first category describes the key attributes of an electronic diary device, a smartphone application, or a web-based module. Five themes specify the intervention.

The first theme, "content," refers to the information in an electronic diary. Smartphone applications and web-based modules consisted of several content types like EMA, reminders, and reward messages [33,40,45,47-49]. This content supports communication between the patient and the health care professional. Long messages are considered too time-consuming to read, and users would therefore skip screens [33,34,42]. Furthermore, users may prefer both cartoons or videos and text [40-42,45,49]. Moreover, diary questions should be tailored to the individual's situation [36,37,42,43]. Users are inconclusive about the scope of the constructs measured; some may prefer an exclusive focus on one topic, whereas others may find that too limited [37,47].

The second theme, "look and feel," refers to the configuration or layout of an electronic diary. The user interface should be both simple and attractive [37,40,45]. However, a balance between attractiveness and user demands is required. Users may prefer a visually appealing user interface with minimal demands on them [45].

The third theme, "functionalities," refers to the activities that a user can perform within an application, ranging from procedures for recording and uploading data to customization of the user interface. Telephone or email reminders, either programmable or automated, notify the user to complete a questionnaire, which increases the completion rate [29-31,33,37,39,42,44,46-48]. Furthermore, manually entering several indicators per day increases participant burden [40-42,45,49]. Moreover, users want to receive motivational feedback about their results via clear graphics and visual displays [33-37,39,42,43,45,47]. Gamification and persuasive techniques can be used to provide motivational feedback to increase completion rates [33,39,40]. Additionally, Tang et al [37] and Triantafyllidis et al [38] identified that technical support and online contact with, for example, a health care professional increase the use of an electronic diary.

The fourth theme, "technological performance," refers to the technological issues that users encounter while using an electronic diary. Users can experience technological issues related to the design concept (eg, navigation problems), the software, or the device (eg, battery attrition). These errors reduce the usability of an electronic diary [30,36,42].

The fifth theme, "preconditions," refers to the conditions that must be fulfilled before a smartphone application or a web-based module can function properly. Burke et al [44] and Tomko et al [49] suggested that users are provided with a compatible mobile device (with sufficient memory, processing speed, and a functioning camera) to overcome the barrier of installing additional hardware or software on the user's device. Moreover, Burke et al [44], Tomko et al [49], and Triantafyllidis et al [38] stated that users need reliable and consistent access to the internet while using the tool. Furthermore, they suggested checking for operating system and other smartphone updates that potentially interfere with the smartphone application of interest [38,44]. The electronic diary should be updated continuously; hence, bandwidth limitations should be taken into account, especially for web-based modules [31]. Automated data transfer to the background server or another device must be seamless for the individual to be able to use the device with minimal effort [36,38]. Depending on the type of data, users highly value data security. They are especially concerned that data would not be shared with health insurers [39].

User Characteristics

The second category describes the characteristics of the individuals who use the electronic diary, in this case, healthy individuals and patients with physical or psychosocial problems. Five themes specify the user characteristics.

The first theme, "sociodemographic information," refers to the characteristics of a population such as gender, age, and marital status. The use of an electronic diary decreases when individuals are older, have a low socioeconomic status, or are unmarried, separated, divorced, or widowed [28,41,49], whereas an increase in the use of these tools is seen when individuals experience high psychological distress [41].

The second theme, "attitudes," refers to the way a user feels and behaves with regard to an electronic diary. Crane et al [45] concluded that users' positive attitudes towards smartphone applications or web-based modules are based on credibility and trustworthiness of the information. Moreover, Tomko et al [49] stated that users may have strong preferences for either electronic or paper diaries.

The third theme, "skills and knowledge," refers to the information that a user has about electronic diaries and the ability to use these tools. Users with no or limited smartphone experience and who experience discomfort with technology will not use electronic diaries adequately. Extra staff is required to train these users [44,49]. Additionally, users who become familiar with self-monitoring or get a sense of mastery over their problems will lose their motivation and consequently stop or reduce their app use [36,39].

The fourth theme, "motivation," refers to the needs, desires, and drives of the individual to use an electronic diary. Naughton et al [42], Anderson et al [39], and Aaron et al [28] stated that missing data are not caused by low motivation, but by discomfort, not having the smartphone at hand, or not wanting to appear rude around others. Social motivation, autonomous motivation, and goal-specific motivation increase the adherence to using electronic diaries [32,37,39]. Furthermore, making users vital partners in the development of an electronic diary keeps them motivated to use these devices [44]. In case of unhealthy behaviors, setting a quit date boosts users' commitment [42].

The fifth theme, "emotional aspects," refers to the feelings that are induced by using an electronic diary. When diary questions are too personal or judgmental, users are less likely to engage with a smartphone application or a web-based module [45]. Furthermore, they want to keep their data private because they are afraid of being judged [45]. However, in the study by Aaron et al [28], emotional aspects were the least mentioned reasons

for missing a questionnaire, although Crane et al [45] found that users feel guilty when diaries are missed.

Process

The third category describes the activities related to the implementation process. One theme specifies the process.

The theme, "training and instructions," refers to how users are guided and instructed to adequately use an electronic diary. Training (eg, face-to-face group kick-off presentation, training session to familiarize with the tool and troubleshoot issues) could result in higher use of these tools [28,30-33,41,43,44,47,49]. Furthermore, users may prefer a troubleshooting guide with step-by-step instructions or continuing technical assistance in case of technological issues from the staff or development team [30,38,44,45,47,49].

Discussion

Principal Findings

This scoping review maps the existing knowledge and gaps concerning factors that influence the use of electronic diaries in health care. Due to technological developments in the last decades, electronic diaries have become increasingly available and popular in research and routine clinical practice. This increased interest is also visible in the large number of articles published between 2000 and 2018. However, only a small number of these articles focused on factors that influence the use of electronic diaries. Additionally, an even smaller number of the selected articles focused on implementing these tools in daily clinical practice.

In this scoping review, 22 articles were selected based on the predefined eligibility criteria. For the categories of intervention, user characteristics, and process of the CFIR [25], 11 themes were identified, whereas no empirical data were found for the 2 other CFIR categories: inner setting and outer setting. The use of an electronic diary is facilitated when it is a visually appealing tool with various content types, including reminders, clear in-app data visualizations tailored to the individual, and minimal user demands to increase the user's engagement. A compatible mobile device with reliable internet access and automated data transfer supports adequate use of an electronic diary. Additionally, the user needs to have smartphone experience, intrinsic motivation, and a clear rationale to monitor one's own behavior. Finally, both theoretical training and practical training are recommended to foster the implementation process. However, the required content and procedures of such training were not described in the included studies.

Based on these results and considering relevant implementation and adoption models, 2 findings attract attention. First, it is remarkable that there were only empirical data about the influence of the characteristics of the electronic diary, the individual, and the implementation process, whereas the CFIR and other implementation frameworks also emphasize the importance of factors related to the organization in which the care is provided or the organizational culture (inner setting) and the competition or the pressure from external partners and the regulations or legislation concerning electronic diaries in clinical practice (outer setting) [25]. Recent research on the implementation of patient-reported outcome measures also highlights the importance of investing sufficient time and resources to support health care professionals [50-54].

Second, the scope of the implementation framework CFIR, used in this review, appears to be wider than adoption models that are traditionally used to evaluate user engagement and continued use of information systems and mobile technologies, like the Technology Acceptance Model [55-59]. The adoption models limit the scope to characteristics of the electronic diary and the individual user, whereas the CFIR also takes into account the process of implementation in daily clinical practice. In this review, the importance of training and instructions was revealed. The importance of hands-on instructions (individual coaching on the job sessions to familiarize with the use of experience-sampling technology in daily clinical practice, using real-world examples) as well as the ability to contact a help desk in case of practical and technological issues was underlined in our previous study as well [27]. Also, regarding the characteristics of the electronic diary, the adoption models have a smaller focus. They only highlight the running software as a contributing factor, while this scoping review identified that the information about and the layout of these diaries, as well as the technological issues and preconditions, also influence their use [55-59]. However, when considering the characteristics of the individual user, this scoping review revealed personal characteristics such as age, along with attitudes, emotions, and behaviors, while adoption models also focus on social influence and self-efficacy as contributing factors [55-59].

Implementation literature emphasizes that attention should be paid to the range of influencing factors to achieve a successful implementation in daily clinical practice [25,50-54]. Consequently, sustainable use of electronic diaries requires that health care organizations or professionals not only direct attention towards software, hardware, and the target population of the tool but also to the economic and political organizational context, the innovation climate in the organization, and the embedding of the tool in routine clinical practice.

Strengths and Limitations

Several limitations have to be kept in mind while interpreting the results of this scoping review. The structured literature search was based on a combination of key words defined by preliminary literature exploration and expert consultation. Despite a broad search approach, it is still possible that articles were missed since the research topic was often not the primary aim of the included studies. This possibly resulted in selection bias. However, the additional hand search minimized this potential shortcoming. It is also worth noting that most of the articles were excluded based on title screening. This can be seen as a limitation, but we think this approach is justifiable in our sensitive search. We performed an iterative screening process that required the researchers to engage in a reflexive way and repeat steps to ensure that the literature was covered in an extensive way. When the relevance of the study was not clear from the title, the abstract was always read. But it is still possible that we missed some articles. Moreover, as an extra check on the 3-step screening process, we read the full texts of a random sample of 50 titles and 50 abstracts. In only 4 articles, we found

information in the results or the discussion related to our scope. Furthermore, as the aim of this scoping review was to map the existing empirical knowledge and identify any gaps about factors that influence the use of electronic diaries in health care, no study quality assessment was performed. Moreover, a scoping review does not endeavor to give a summary of the existing literature or compare results (in contrast to a systematic review of, for example, randomized controlled trials on efficacy). Therefore, we did not intend to draw firm conclusions regarding useful and effective features of electronic diaries based on quantified outcomes. We provide, to our knowledge, a first overview of the factors that influence the use of electronic diaries in health care. Future research with longitudinal or mixed methods study designs should focus on the causal relationships between the influencing factors and the use of electronic diaries in health care in order to get a deeper understanding of the causality. Also, a quite diverse sample of studies was included. However, we are convinced that we have achieved the scope of interest of this scoping review. We looked in more detail at similarities and differences in the results of the included studies, based on the purpose of use (monitoring versus intervention), target population (healthy individuals versus patients), setting, study aims, and design (feasibility versus usability versus development). However, we concluded that this synthesis cannot be performed based on the results of the information found in this scoping review. More research is needed in this field. Additionally, the structured literature search was restricted to peer-reviewed databases and so, empirical research. Book chapters and grey literature were not included, which means that additional empirical data can be lacking. This scoping review has several methodological strengths as well. First, a systematic approach was used based on the methodological framework by Arksey and O'Malley [24]. The interprofessional nature of the research team extended the scope of this review, and the consultation of 2 experts in the field validated the search terms. Furthermore, the 3-step screening process was consistently performed by 2 researchers. Second, the thematic analysis organized according to an implementation research perspective led to a synthesis contributing to future understanding of the implementation of electronic diaries in health care.

Conclusion

This scoping review demonstrates that the use of electronic diaries may be influenced by characteristics of the electronic diary, the individual user, and the implementation process. However, the number of empirical studies on the topic was limited. Studies that take into account the setting in which to implement the diaries, such as the organizational context, the implementation climate, and available organizational resources, were lacking. Future research should focus on these factors and on the causal relationships between the different factors to investigate the continued use of these innovative tools.

Acknowledgments

We thank Gaston Jamin and François Engelen for sharing their ideas about key authors and key publications on the topic of interest.

This project was supported by the Brightlands Innovation Program LIME (Limburg Measures; limeconnect.nl), a program that facilitates smarter measurement methods and more efficient data collection for better care and health; the Province of Limburg; Zuyd University of Applied Sciences; and Maastricht University, the Netherlands.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1. Patterns of publications.
[DOCX File , 13 KB - mhealth_v9i6e19536_app1.docx]

References

1. Shindorf Z. Exploring Barriers to the Generalization of Social Skills Interventions for Children Diagnosed with ASD: A Qualitative Analysis of "Youth Engagement Through Intervention". University of Montana. 2016. URL: https://scholarworks.umt.edu/etd/10911/ [accessed 2019-09-10]
2. Smith PC, Araya-Guerra R, Bublitz C, Parnes B, Dickinson LM, Van Vorst R, et al. Missing clinical information during primary care visits. JAMA 2005 Feb 02;293(5):565-571. [doi: 10.1001/jama.293.5.565] [Medline: 15687311]
3. Higgs J, Jones M, Loftus S, Christensen N, editors. Clinical reasoning in the health professions, fourth edition. Amsterdam, Netherlands: Elsevier; 2018.
4. van Os J, Delespaul P, Wigman J, Myin-Germeys I, Wichers M. Beyond DSM and ICD: introducing "precision diagnosis" for psychiatry using momentary assessment technology. World Psychiatry 2013 Jun;12(2):113-117 [FREE Full text] [doi: 10.1002/wps.20046] [Medline: 23737412]
5. van Os J, Lataster T, Delespaul P, Wichers M, Myin-Germeys I. Evidence that a psychopathology interactome has diagnostic value, predicting clinical needs: an experience sampling study. PLoS One 2014;9(1):e86652 [FREE Full text] [doi: 10.1371/journal.pone.0086652] [Medline: 24466189]

6. National Research Council. Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. In: National Academies 030907. Washington, DC: The National Academies Press; 1984.
7. Stone A, Bachrach C, Jobe J, Kurtzman H, Cain V, editors. The science of self-report: Implications for research and practice. New York, NY: Psychology Press; 2000.
8. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ 2002 May 18;324(7347):1193-1194 [FREE Full text] [doi: 10.1136/bmj.324.7347.1193] [Medline: 12016186]
9. Byrom B, Tiplady B, editors. ePRO: electronic solutions for patient-reported data. Abingdon-on-Thames, Oxfordshire, England: Routledge; 2010.
10. Dale O, Hagen KB. Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. J Clin Epidemiol 2007 Jan;60(1):8-17. [doi: 10.1016/j.jclinepi.2006.04.005] [Medline: 17161749]
11. Milligan C, Bingley A, Gatrell A. Digging deep: using diary techniques to explore the place of health and well-being amongst older people. Soc Sci Med 2005 Nov;61(9):1882-1892. [doi: 10.1016/j.socscimed.2005.04.002] [Medline: 15913863]
12. Ladores S, Bray LA. Electronic Diaries in Healthcare: A Review of Literature. JNPARR 2009:13-21. [doi: 10.13178/jnparr.2017.0702.0705]
13. Green AS, Rafaeli E, Bolger N, Shrout PE, Reis HT. Paper or plastic? Data equivalence in paper and electronic diaries. Psychol Methods 2006 Mar;11(1):87-105. [doi: 10.1037/1082-989X.11.1.87] [Medline: 16594769]
14. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials 2003 Apr;24(2):182-199. [doi: 10.1016/s0197-2456(02)00320-3] [Medline: 12689739]
15. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54:579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]
16. Drummond HE, Ghosh S, Ferguson A, Brackenridge D, Tiplady B. Electronic quality of life questionnaires: a comparison of pen-based electronic questionnaires with conventional paper in a gastrointestinal study. Qual Life Res 1995 Feb;4(1):21-26. [doi: 10.1007/BF00434379] [Medline: 7711686]
17. Finkelstein J, Cabrera MR, Hripcsak G. Internet-based home asthma telemonitoring: can patients handle the technology? Chest 2000 Jan;117(1):148-155. [doi: 10.1378/chest.117.1.148] [Medline: 10631213]
18. Hufford M, Shields A. Electronic diaries: Applications and What Works in the Field. Applied Clinical Trials 2002:46-59.
19. Totterdell P, Folkard S. In situ repeated measures of affect and cognitive performance facilitated by use of a hand-held computer. Behavior Research Methods, Instruments, & Computers 1992 Dec;24(4):545-553. [doi: 10.3758/bf03203603]
20. Kamarck TW, Shiffman SM, Smithline L, Goodie JL, Paty JA, Gnys M, et al. Effects of task strain, social conflict, and emotional activation on ambulatory cardiovascular activity: daily life consequences of recurring stress in a multiethnic adult sample. Health Psychol 1998 Jan;17(1):17-29. [doi: 10.1037//0278-6133.17.1.17] [Medline: 9459066]
21. Hsiao C, Chang J, Tang K. Exploring the influential factors in continuance usage of mobile social Apps: Satisfaction, habit, and customer value perspectives. Telematics and Informatics 2016 May;33(2):342-355. [doi: 10.1016/j.tele.2015.08.014]
22. Kang S. Factors influencing intention of mobile application use. IJMC 2014;12(4):360. [doi: 10.1504/ijmc.2014.063653]
23. Wang BR, Park J, Chung K, Choi IY. Influential Factors of Smart Health Users according to Usage Experience and Intention to Use. Wireless Pers Commun 2014 Apr 24;79(4):2671-2683. [doi: 10.1007/s11277-014-1769-0]
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
25. 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] [Medline: 19664226]
26. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
27. Daniëls NEM, Hochstenbach LMJ, van Bokhoven MA, Beurskens AJHM, Delespaul PAEG. Implementing Experience Sampling Technology for Functional Analysis in Family Medicine - A Design Thinking Approach. Front Psychol 2019 Dec 11;10:2782 [FREE Full text] [doi: 10.3389/fpsyg.2019.02782] [Medline: 31920830]
28. Aaron LA, Mancl L, Turner JA, Sawchuk CN, Klein KM. Reasons for missing interviews in the daily electronic assessment of pain, mood, and stress. Pain 2004 Jun;109(3):389-398. [doi: 10.1016/j.pain.2004.02.014] [Medline: 15157700]
29. Litcher-Kelly L, Kellerman Q, Hanauer SB, Stone AA. Feasibility and utility of an electronic diary to assess self-report symptoms in patients with inflammatory bowel disease. Ann Behav Med 2007 Apr;33(2):207-212. [doi: 10.1007/BF02879902] [Medline: 17447873]
30. Welch J, Dowell S, Johnson C. Feasibility of using a personal digital assistant to self-monitor diet and fluid intake: a pilot study. Nephrol Nurs J 2007;34(1):43-8; quiz 49. [Medline: 17345691]
31. Stevens VJ, Funk KL, Brantley PJ, Erlinger TP, Myers VH, Champagne CM, et al. Design and implementation of an interactive website to support long-term maintenance of weight loss. J Med Internet Res 2008 Jan 25;10(1):e1 [FREE Full text] [doi: 10.2196/jmir.931] [Medline: 18244892]
32. Webber KH, Tate DF, Ward DS, Bowling JM. Motivation and its relationship to adherence to self-monitoring and weight loss in a 16-week Internet behavioral weight loss intervention. J Nutr Educ Behav 2010;42(3):161-167. [doi: 10.1016/j.jneb.2009.03.001] [Medline: 20138583]

33. Ahtinen A, Mattila E, Välkkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. JMIR Mhealth Uhealth 2013 Jul 10;1(2):e11 [FREE Full text] [doi: 10.2196/mhealth.2596] [Medline: 25100683]
34. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013 Dec;36(4):289-296 [FREE Full text] [doi: 10.1037/prj0000019] [Medline: 24015913]
35. Ma J, Xiao L, Blonstein AC. Measurement of self-monitoring web technology acceptance and use in an e-health weight-loss trial. Telemed J E Health 2013 Oct;19(10):739-745 [FREE Full text] [doi: 10.1089/tmj.2013.0009] [Medline: 23952787]
36. Tatara N, Arsand E, Skrøvseth SO, Hartvigsen G. Long-term engagement with a mobile self-management system for people with type 2 diabetes. JMIR Mhealth Uhealth 2013 Mar 27;1(1):e1 [FREE Full text] [doi: 10.2196/mhealth.2432] [Medline: 25100649]
37. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
38. 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] [Medline: 26037921]
39. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
40. Batink T, Bakker J, Vaessen T, Kasanova Z, Collip D, van Os J, et al. Acceptance and Commitment Therapy in Daily Life Training: A Feasibility Study of an mHealth Intervention. JMIR Mhealth Uhealth 2016 Sep 15;4(3):e103 [FREE Full text] [doi: 10.2196/mhealth.5437] [Medline: 27634747]
41. Jiang Y, Sereika SM, Dabbs AD, Handler SM, Schlenk EA. Acceptance and Use of Mobile Technology for Health Self-Monitoring in Lung Transplant Recipients during the First Year Post-Transplantation. Appl Clin Inform 2016;7(2):430-445 [FREE Full text] [doi: 10.4338/ACI-2015-12-RA-0170] [Medline: 27437052]
42. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A Context-Sensing Mobile Phone App (Q Sense) for Smoking Cessation: A Mixed-Methods Study. JMIR Mhealth Uhealth 2016 Sep 16;4(3):e106 [FREE Full text] [doi: 10.2196/mhealth.5787] [Medline: 27637405]
43. Timmerman JG, Tönis TM, Dekker-van Weering MGH, Stuiver MM, Wouters MWJM, van Harten WH, et al. Co-creation of an ICT-supported cancer rehabilitation application for resected lung cancer survivors: design and evaluation. BMC Health Serv Res 2016 Apr 27;16:155 [FREE Full text] [doi: 10.1186/s12913-016-1385-7] [Medline: 27121869]
44. 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] [Medline: 28298264]
45. Crane D, Garnett C, Brown J, West R, Michie S. Factors Influencing Usability of a Smartphone App to Reduce Excessive Alcohol Consumption: Think Aloud and Interview Studies. Front Public Health 2017;5:39 [FREE Full text] [doi: 10.3389/fpubh.2017.00039] [Medline: 28421175]
46. Freyne J, Yin J, Brindal E, Hendrie GA, Berkovsky S, Noakes M. Push Notifications in Diet Apps: Influencing Engagement Times and Tasks. International Journal of Human–Computer Interaction 2017 Jan 11;33(10):833-845. [doi: 10.1080/10447318.2017.1278896]
47. Kreyenbuhl J, Record EJ, Himelhoch S, Charlotte M, Palmer-Bacon J, Dixon LB, et al. Development and Feasibility Testing of a Smartphone Intervention to Improve Adherence to Antipsychotic Medications. Clin Schizophr Relat Psychoses 2019;12(4):152-167 [FREE Full text] [doi: 10.3371/CSRP.KRRE.070816] [Medline: 27454213]
48. Liu S, Willoughby JF. Do Fitness Apps Need Text Reminders? An Experiment Testing Goal-Setting Text Message Reminders to Promote Self-Monitoring. J Health Commun 2018;23(4):379-386. [doi: 10.1080/10810730.2018.1455768] [Medline: 29601270]
49. Tomko RL, Gray KM, Oppenheimer SR, Wahlquist AE, McClure EA. Using REDCap for ambulatory assessment: Implementation in a clinical trial for smoking cessation to augment in-person data collection. Am J Drug Alcohol Abuse 2018;45(1):26-41 [FREE Full text] [doi: 10.1080/00952990.2018.1437445] [Medline: 29513609]
50. Antunes B, Harding R, Higginson IJ, EUROIMPACT. Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. Palliat Med 2014 Feb;28(2):158-175 [FREE Full text] [doi: 10.1177/0269216313491619] [Medline: 23801463]
51. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf 2014 Jun;23(6):508-518. [doi: 10.1136/bmjqs-2013-002524] [Medline: 24505110]
52. Duncan EA, Murray J. The barriers and facilitators to routine outcome measurement by allied health professionals in practice: a systematic review. BMC Health Serv Res 2012 May 22;12(1):96 [FREE Full text] [doi: 10.1186/1472-6963-12-96] [Medline: 22506982]

53. Greenhalgh J, Dalkin S, Gooding K, Gibbons E, Wright J, Meads D, et al. Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. Health Serv Deliv Res 2017 Jan;5(2):1-280. [doi: 10.3310/hsdr05020] [Medline: 28121094]
54. Howell D, Molloy S, Wilkinson K, Green E, Orchard K, Wang K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. Ann Oncol 2015 Sep;26(9):1846-1858 [FREE Full text] [doi: 10.1093/annonc/mdv181] [Medline: 25888610]
55. Rogers EM. Diffusion of Innovations: modifications of a model for telecommunications. In: Stoetzer MW, Mahler A, editors. Die diffusion von innovationen in der telekommunikation. Schriftenreihe des Wissenschaftlichen Instituts für Kommunikationsdienste, vol 17. Berlin, Heidelberg: Springer; 1995.
56. Ajzen I. From Intentions to Actions: A Theory of Planned Behavior. In: Kuhl J, Beckmann J, editors. Action Control. SSSP Springer Series in Social Psychology. Berlin, Heidelberg: Springer; 1985:11-39.
57. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
58. Bandura A. Self-Efficacy: The exercise of control. New York, NY: Worth Publishers; 1997.
59. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
EMA: ecological momentary assessment

Edited by L Buis; submitted 22.04.20; peer-reviewed by A Triantafyllidis, D Salvi, M Dunbar, J Willoughby; comments to author 04.09.20; revised version received 30.10.20; accepted 25.02.21; published 01.06.21.

Please cite as:
Daniëls NEM, Hochstenbach LMJ, van Zelst C, van Bokhoven MA, Delespaul PAEG, Beurskens AJHM
Factors That Influence the Use of Electronic Diaries in Health Care: Scoping Review
JMIR Mhealth Uhealth 2021;9(6):e19536
URL: https://mhealth.jmir.org/2021/6/e19536
doi:10.2196/19536
PMID:34061036

Review

Older Adults' Experiences With Using Wearable Devices: Qualitative Systematic Review and Meta-synthesis

Kevin Moore[1], MPH; Emma O'Shea[1], PhD; Lorna Kenny[1], MSc; John Barton[2], MEng; Salvatore Tedesco[2], MSc; Marco Sica[2], MSc; Colum Crowe[2], MSc; Antti Alamäki[3], PhD; Joan Condell[4], PhD; Anna Nordström[5,6], PhD; Suzanne Timmons[1], MD

[1]Centre for Gerontology and Rehabilitation, University College Cork, Cork, Ireland

[2]Tyndall Institute, Cork, Ireland

[3]Research, Development and Innovation Activities & Physiotherapy Education, Karelia University of Applied Sciences, Karelia, Finland

[4]School of Computing, Engineering and Intelligent Systems, Faculty of Computing, Engineering and the Built Environment, Ulster University, Coleraine, United Kingdom

[5]Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

[6]School of Sport Sciences, The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Kevin Moore, MPH
Centre for Gerontology and Rehabilitation
University College Cork
The Bungalow, Block 13
St Finbarrs Hospital, Douglas Road
Cork, T12 Y319
Ireland
Phone: 353 857070042
Email: kevin.moore@ucc.ie

Abstract

Background: Older adults may use wearable devices for various reasons, ranging from monitoring clinically relevant health metrics or detecting falls to monitoring physical activity. Little is known about how this population engages with wearable devices, and no qualitative synthesis exists to describe their shared experiences with long-term use.

Objective: This study aims to synthesize qualitative studies of user experience after a multi-day trial with a wearable device to understand user experience and the factors that contribute to the acceptance and use of wearable devices.

Methods: We conducted a systematic search in CINAHL, APA PsycINFO, PubMed, and Embase (2015-2020; English) with fixed search terms relating to *older adults* and *wearable devices*. A meta-synthesis methodology was used. We extracted themes from primary studies, identified key concepts, and applied reciprocal and refutational translation techniques; findings were synthesized into third-order interpretations, and finally, a "line-of-argument" was developed. Our overall goal was theory development, higher-level abstraction, and generalizability for making this group of qualitative findings more accessible.

Results: In total, we reviewed 20 papers; 2 evaluated fall detection devices, 1 tested an ankle-worn step counter, and the remaining 17 tested activity trackers. The duration of wearing ranged from 3 days to 24 months. The views of 349 participants (age: range 51-94 years) were synthesized. Four key concepts were identified and outlined: motivation for device use, user characteristics (openness to engage and functional ability), integration into daily life, and device features. Motivation for device use is intrinsic and extrinsic, encompassing many aspects of the user experience, and appears to be as, if not more, important than the actual device features. To overcome usability barriers, an older adult must be motivated by the useful purpose of the device. A device that serves its intended purpose adds value to the user's life. The user's needs and the support structure around the device—aspects that are often overlooked—seem to play a crucial role in long-term adoption. Our "line-of-argument" model describes how motivation, ease of use, and device purpose determine whether a device is perceived to add value to the user's life, which subsequently predicts whether the device will be integrated into the user's life.

Conclusions: The added value of a wearable device is the resulting balance of motivators (or lack thereof), device features (and their accuracy), ease of use, device purpose, and user experience. The added value contributes to the successful integration of the

device into the daily life of the user. Useful device features alone do not lead to continued use. A support structure should be placed around the user to foster motivation, encourage peer engagement, and adapt to the user's preferences.

(JMIR Mhealth Uhealth 2021;9(6):e23832) doi:10.2196/23832

KEYWORDS

wearable device; older adult; digital health; meta-synthesis; qualitative review; acceptance; adherence; mobile phone

Introduction

Background

Wearable health monitoring devices have seen a rapid rise in capability and popularity over the last two decades. These small wireless devices can monitor movements, improve physical activity, and facilitate ageing-in-place. Wearable devices temporarily and noninvasively attach to a person without hindering their movement and are often intended to be worn continuously. Examples include activity trackers (eg, Fitbit and smartwatches), fall detection devices, electromyography patches, and smart clothing.

Although older adults are not core consumers of wearable devices, their use of digital health technologies is increasing [1] in tandem with the expanding technological capabilities of wearable devices. Wearable devices can support "active ageing," the process of enhancing quality of life as people age [2]. Technology creates an enabling environment that restores function and expands the participation of older adults in their health. Remote monitoring using wearable devices can aid independence and encourage older adults to manage stable chronic conditions by themselves. Clinicians can track patients' health status remotely and communicate via video-based consultations [3]. Current wearable devices possess the ability to monitor a number of health metrics, including heart rate, blood oxygen levels, body temperature, physical activity, sleep, and blood pressure [4]. The older adult population is vulnerable to changes in their health conditions and may be burdened by frequent clinical visits. Wearable devices are well suited for monitoring older adults because they convey up-to-date health information and track health metrics over time. Wearable devices are intended to be worn continuously. For example, fall detection devices are worn all day, as falls occur unexpectedly. As these devices are used frequently, it is important to understand the barriers to acceptance and adherence. Factors such as trust, functionality, added value, ease of use, cost, stigma, and fear of dependence are examples of barriers to adoption [5].

Researchers have used a variety of methods to collect information from older adults regarding the acceptability of wearable devices: surveys [6,7], wear time [1], diaries [8,9], interviews [9], and focus groups [10,11]. Some studies collected information about general preferences regarding device design [11]; others allowed participants to interact with several devices before asking about preferred design features [12], in which participants used a wearable device for multiple days and then provided feedback [13-15].

Objectives

Qualitative research methods are well suited to examine the user experience and may offer explanations for unexpected or anomalous findings in quantitative data [16] or uncover usability barriers that quantitative approaches often miss. Systematic reviews that combine the findings of multiple qualitative studies can identify common factors among studies and generalize their findings. No qualitative systematic review exists on older adults' experiences of using any form of wearable device. Although each user experience is unique, a synthesis of studies may lead to a richer understanding of the integration of devices into the lives of users. Our objective is to better understand these experiences to inform future research efforts and to inspire device design to ensure a successful user experience.

We aim to apply a qualitative meta-synthesis process to the available qualitative data on older adults' experiences with using wearable devices. Meta-synthesis is a form of interpretive synthesis that can be used in the review and evaluation of qualitative research studies. Our meta-synthesis is based on the principles of meta-ethnography [17], a method designed by Noblit and Hare [17] to synthesize ethnographic studies. A meta-synthesis differs from traditional meta-ethnography in that it allows for a variety of data analysis techniques besides ethnography (eg, phenomenology and grounded theory) to be synthesized together. It is an inductive method that compares, translates, and integrates concepts across studies, while also preserving the context of the primary data. This meta-synthesis is more in-depth than previous systematic reviews on wearable devices that summarized a group of studies [4,5,18-23], analyzed a series of trials [24], or reviewed the state of the art [25,26]. Our overall goal is theory development, higher-level abstraction, and generalizability for making this group of qualitative findings more accessible [27].

Our overarching research question is "What is the experience of older adults who took part in multi-day wearable device trials and what factors contribute to acceptance and use?"

Methods

Search Strategy

The inclusion and exclusion criteria (Textbox 1) were designed to accommodate various study designs as long as they contained the following core structure: an older person using a wearable device for multiple days and the subsequent qualitative analysis of that participant's experience in relation to the wearable device. We created an inclusive search strategy to locate studies that used unusual jargons or unconventional study designs (Textbox 2). We also reviewed the search terms used in other systematic reviews in this area, consulted with colleagues in clinical and technical expertise in the area, and piloted various

combinations of search terms to assess the sensitivity of the terms. KM searched four databases—CINAHL, APA PsycINFO, PubMed, and Embase—from January 2015 to January 2020 for studies published in English. Our date range intentionally excluded older wearable devices to minimize the differences between the capabilities of the devices used in the studies. Insights from before this period were recorded in a systematic review by Bergmann and McGregor [28]. This search was supplemented by searching Google Scholar, forward and backward searches of citation lists, and the publication lists of prominent researchers in the field.

Textbox 1. Inclusion and exclusion criteria guiding study selection.

Inclusion criteria

- Peer-reviewed studies (in English)
- Published between January 2015 and January 2020
- Experiences of older adults using wearable devices
- Using a defined qualitative approach
- Presenting distinct qualitative data and results
- Qualitative data collected after the multi-day trial

Exclusion criteria

- Studies not in English or outside the time frame
- Not focused on older populations
- No primary qualitative data presented
- No continuous, multi-day trial component

Textbox 2. Search terms used in the search strings.

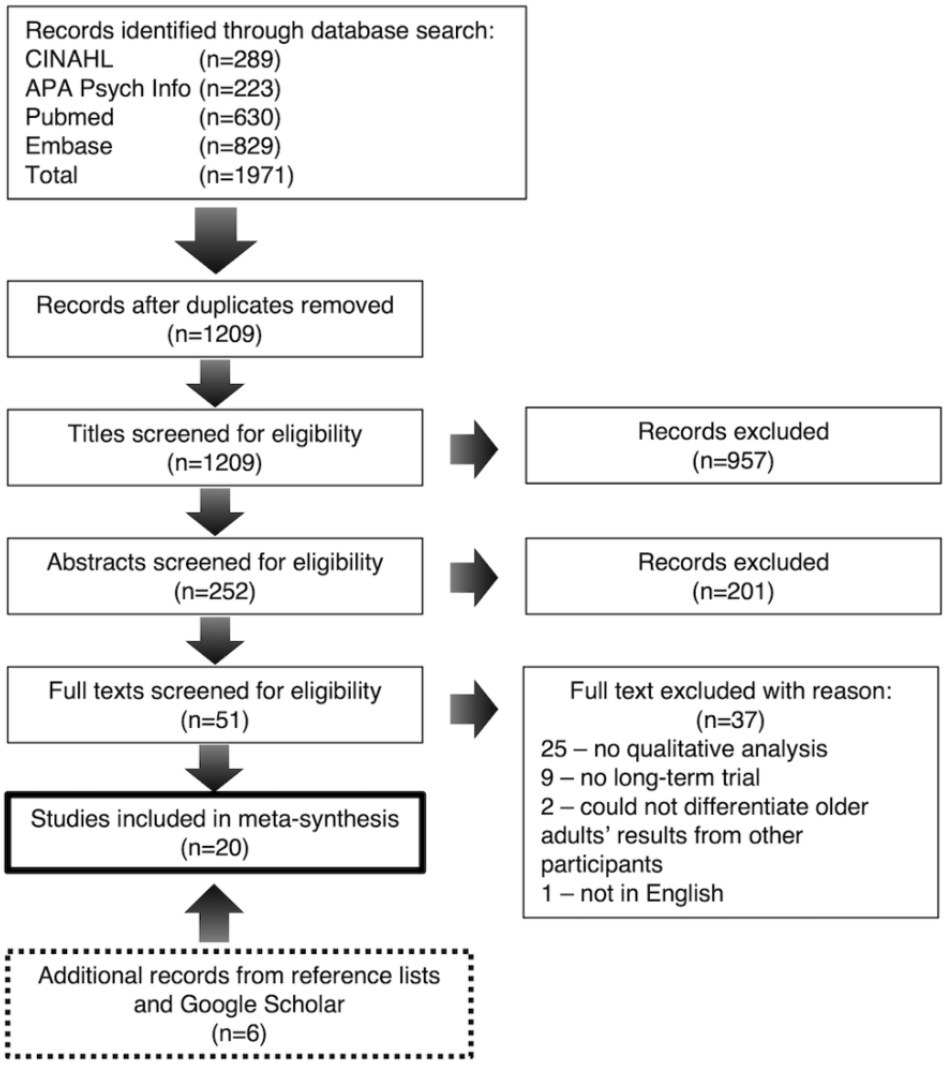
Search terms

- (“wearable technology” OR “wearable sensor” OR “inertia sensor” OR “wireless sensor” OR accelerometer OR “Micro-Electrical-Mechanical-System” OR Actigraphy OR “inertial measurement unit” OR “motion monitor” OR “movement sensor” OR “wearable interface*” OR “body worn” OR wearable OR “wireless monitoring system” OR “activity tracker*” OR “activity-tracking” OR “activity sensor*” OR “activity assessment*” OR “fall detection” OR “wireless sensor networks”) AND (“user preference*” OR “user experience*” OR “user needs” OR preference* OR “patient centered” OR qualitative OR “focus group” OR perception* OR understanding OR acceptance OR adoption OR usability OR perspective) AND (“older adults” OR older OR ageing OR Parkinson’s OR Alzheimer’s OR Dementia OR stroke OR chronic) NOT (invasive OR implant*)

Data Collection

KM exported the search results, removed duplicates, screened all titles and abstracts for inclusion, and reviewed eligible full-text articles against the inclusion or exclusion criteria (Textbox 1). LK screened a random sample of 200 abstracts (and full texts where the abstract indicated potential for inclusion) and confirmed the consistent application of the inclusion and exclusion criteria. A senior researcher, ST, provided guidance when eligibility based on a full-text review was unclear. Reasons for exclusion were recorded for all excluded studies (Figure 1).

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of study selection process.



Data Extraction and Quality Appraisal

Data items extracted included information about the publication (date, authors, and study aims), study process (design, methods, and analysis), participant characteristics, device types and features, trial length, and relevant primary qualitative data (themes and quotations).

Although not required in a meta-ethnography, we assessed study quality to facilitate the critical reading of each study to gauge its potential contribution to the analysis (see Table S2 in Multimedia Appendix 1 [9,13-15,29-44] for the checklist). We used the Evaluation Tool for Qualitative Studies (ETQS) [45], as this provides detailed instructions on applying the evaluation criteria, unlike the CASP (Critical Appraisals Skills Programme) tool [46,47]. The ETQS guides the appraisal of the phenomenon studied and context issues; ethics; data collection, analysis, and researcher bias; and policy and practice implications [45]. Two authors (KM and LK) independently conducted the quality assessment. No studies were excluded based on ETQS results, as this often reflects the level of reporting transparency, rather than the actual research processes used [16,48].

Data Analysis and Synthesis

The analysis (Textbox 3) was guided by a meta-ethnographic approach [17,48,49]. Initially, papers were read and reread to familiarize researchers with the study context, design, and findings. Individual themes were extracted and recorded using separate index cards. As the studies were methodologically heterogeneous, we preserved the authors' original themes (and wording), where possible. Where appropriate, we extracted additional themes from the "discussion" and "conclusion" sections. Where studies were highly descriptive or simply listed participants' quotes, we coded the "results" and "discussion" sections and generated themes from the presented data, noting that specific quotes, without the full conversation context, were challenging to code. Studies that presented minimal or overly descriptive results were mainly used to support or refute the themes identified in high-quality studies.

Textbox 3. Key steps involved in the synthesis and adapted from Noblit and Hare.

Key Steps

- Identifying knowledge gaps and the literature available for a synthesis and developing research questions
- Defining the focus of the synthesis, locating relevant studies, and assessing the quality of the included studies
- Active reading of the studies to understand context and to extract relevant data
- Themes and concepts were identified in the "results" and "discussion" sections; authors' interpretations were retained where possible; descriptive studies (without author-generated themes) were coded; and each theme was transferred to an index card along with contextual information, a narrative summary, and device characteristics.
- Index cards (each containing an extracted theme) were juxtaposed and grouped into general categories, categories were refined and subcategories emerged, and key concepts were identified.
- Returning to each study and comparing with the generated categories, using the context provided by the authors to re-evaluate the category placement of each index card, generating new index cards when the existing index cards do not represent the totality of the study results, and generating new subcategories and condensing others to better describe the results of the studies
- Compiling the participant raw data, index cards, and categories to produce overarching concepts that describe the results of the translation process
- Development of a line-of-argument synthesis and conceptual model

The themed index cards were physically juxtaposed and grouped into categories based on patterns of meaning, as related to the research question. Categories and subcategories were refined iteratively through constant comparison within and across studies. Each category was compared against each original study using (1) reciprocal translation (recognizing reoccurring themes or concepts across studies) and (2) refutational translation (recognizing dissimilar themes or concepts across studies, not explained by contextual factors). When all the data were collated and interpreted, several key concepts were defined (third order) and synthesized to develop an integrative "line-of-argument."

We tracked the preferred and disliked device features throughout the analysis process. Where relevant, we used specific device features to support the key concepts. We summarized the preferred and disliked features, but no frequency analysis was performed because each study used different devices and not all studies included participant feedback on device features.

We reported our results in line with the eMERGe guidance, which has been described for use by researchers conducting meta-ethnography [50]. The search strategy results are presented in a Preferred Reporting Items for Systematic Reviews and Meta-analyses flow diagram (Figure 1).

Results

Overview

The database search returned 1971 results (Figure 1). After title or abstract screening and duplicate removal, 51 full-text records were reviewed, and 14 were considered eligible for inclusion. Backward or forward searching uncovered 6 additional eligible records to reach a total of 20 records for the synthesis.

Characteristics of Included Studies

Of the 20 included records, 2 evaluated fall detection devices [14,29], 1 tested an ankle-worn step counter [30], and the remaining 17 examined wrist-worn activity trackers. The duration of use ranged from 3 days to 24 months. In some studies, users completed multiple trials with different devices [9,31-33]. In others, participants were randomly assigned to one of several devices [34,35]. The views of 349 participants (age: range 51-94 years) were synthesized, including those with previous breast cancer [32], obesity [36,37], resolving heart failure [30,37], Parkinson disease [38], dementia [15,39], and walking aids [9,29] and those who were fully independent and healthy [9,13,14,29,31-35,37,40-44]. Table 1 summarizes the results of the data extraction process.

Table 1. Articles included in the meta-synthesis and quality appraisal scores using the Evaluation Tool for Qualitative Studies.

Study	Method	Participants	Device	Trial duration	ETQS[a] [45]
Abouzahra and Ghasemaghaei [13]	• Interview pre- and posttrial[b] • Device data	• 44 participants • Aged 65-75 years	• Fitbit: AT[c]+SP[d]—wrist-worn	1 week	6
Batsis [36]	• Surveys • Interview[b]	• 8 participants • Aged 65-80 years • Rural; obese	• Fitbit: PD[e]+SP—waist clip	4 weeks	5
Demiris et al [14]	• Interview (×2)[b] • Fall or device log	• 18 participants • Aged ≥62 years • Slight fall risk	• FDD[f]—clip or lanyard	4 months	8
Ehn et al [9]	• Interview[b] • Follow-up meeting • Activity diary	• 8 participants • Aged 75-90 years	• Withings Activité Pop: AT+SP—wrist-worn • Jawbone UP3: AT+SP—wrist-worn	9-10 days for each device	10
Farina et al [15]	• Device diary • Questionnaire • Dyadic interview[b]	• 26 participants • Aged 65-90 years • Alzheimer and dementia	• GENEactiv Original: AT—wrist-worn	1 month	8
Fausset et al [34]	• Questionnaire—interviews pre- or posttrial[g] • Daily diary	• 8 participants • Aged 61-69 years	• Striiv: PD—clip • Fitbit: PD—clip • Nike+FuelBand: AT—wrist-worn • MyFitnessPal: web-based	Randomly assigned one device for 2 weeks	5
Floegel et al [30]	• Interview[b]	• 27 participants • Aged 62-90 years • Heart failure requiring hospitalization	• Tractivity: AT+SP—ankle-worn	1 month	6
Hermanns et al [38]	• Surveys[g] • Interview[b]	• 5 participants • Aged 65-81 years • Stage I-IV Parkinson disease	• Fitbit: AT+tablet—wrist-worn	12 weeks	8
Kononova et al [37]	• Focus group[a]	• 48 (nonusers, short-term, former, and long-term users) • Aged 65-94 years	• Garmin Vivofit 2: AT—wrist-worn	2-4 weeks	9
Lee et al [40]	• Adoption and usability surveys • Biweekly interviews[b]	• 17 participants • Aged 65-85 years	• Nokia Go: AT—wrist-worn	14 weeks	8
Mercer et al [31]	• Questionnaire • Focus groups[b]	• 32 participants • Aged 52-84 years	• Fitbit Zip: PD—clip • Jawbone Up 24: AT—wrist-worn • Misfit Shine: AT—wrist-worn or clip • Withings Pulse: AT—wrist-worn • PD—clip	5 devices, each for ≥3 days (≥15 days total)	8

Study	Method	Participants	Device	Trial duration	ETQS[a] [45]
Nguyen et al [32]	• Focus groups[b]	• 14 participants • Aged 51-64 years	• Fitbit One: PD—clip • Jawbone Up 24: AT—wrist-worn • Garmin Vivofit 2: AT—wrist-worn • Garmin Vivosmart: AT—wrist-worn • Garmin Vivoactive: AT—wrist-worn • Polar A300: AT—wrist-worn	Assigned 2 devices; 2 weeks per device, 4 weeks total	8
Preusse et al [35]	• Questionnaire; interview[h]	• 16 participants • Aged 65-73 years	• MyFitnessPal: web-based • Fitbit One: AT—wrist-worn+web-based	28 days	7
Puri et al [33]	• Questionnaire • Interview with sample[i]	• 20 participants • Aged 55-84 years	• Microsoft Band: AT—wrist-worn • Mi Band: AT—wrist-worn	Each device for 3 weeks, 6 weeks total	9
Rosales et al [41]	• Interviews (×2)[h]	• 5 participants • Aged ≥65 years • Smartphone users	• Moto G 360: SW—wrist-worn	12-24 months	6
Schlomann et al [43]	• Group discussion[b]	• 6 participants • Aged 67-78 years	• AT+SP—wrist-worn	1 month	6
Schlomann [42]	• Interviews (×2)[b,h]	• 6 participants • Aged 60-78 years • Smartphone users	• ViFit: AT—wrist-worn	1 year	6
Thilo et al [29]	• Daily diary • Focus group[h]	• 15 participants • Aged 75-92 years • History of falls	• FDD—torso patch	9 days	8
Thorpe et al [39]	• Interview[h] • Sensor data	• 6 participants • Aged 65-78 years • Dementia	• Sony SmartWatch 3: SW+SP—wrist-worn	9 weeks	6
Zhou et al [44]	• Questionnaire • Interview[b]	• 20 participants • Aged 58-68 years	• 37 Degree Technology: AT+SP—wrist-worn	3 months	6

[a]ETQS: Evaluation Tool for Qualitative Studies; maximum score is 10.

[b]Inductive analysis.

[c]AT: activity tracker.

[d]SP: smartphone.

[e]PD: pedometer.

[f]FDD: fall detection device.

[g]Descriptive analysis.

[h]Deductive analysis.

[i]Directed content analysis.

Translation

This study found four key concepts, comprising 12 subthemes that characterize the collective experience of trial participants (Textbox 4).

Textbox 4. Results of the reciprocal and refutation translation process.

Category 1: Openness to engage and functional ability of the user

- Age-related physiology and comorbidities [14,32]
  - Physical limitations such as hand dexterity
  - Slower processing speeds in time of need
  - Inactive lifestyle does not warrant activity tracker
- Sense of independence [14,15,29,33]
  - Confidence in abilities to remember procedures
  - Change in routine (battery life; attached to phone)
  - Subjective norm, not burden on family
  - Access to instructions and training
- Exploration and use of device features [9,32,43]
  - Interest in diverse features and uses
  - Confidence to explore means maximized benefits
  - Technology experience means ability to troubleshoot
  - Instructions to overcome hurdles
- Self-efficacy for technology [9,13,31]
  - Skill to control and manipulate technology
  - Perception of one's own ability to use technology
  - Insecurities of using the system reduced usage

Category 2: Motivation for device use

- Awareness of physical activity levels [9,31,32,37,38]
  - Real versus perceived activity levels
  - Awareness is not the same as motivation
  - Awareness is a catalyst, not a creator of motivation
- Internal influences [9,13,37,39]
  - Intrinsic motivation required for behavior change
  - Achieving personal goals is satisfying
  - Desire to improve health and fitness
  - Expectation-confirmation theory; if the device meets the user's expectation, they may be more likely to adopt the device
- Quantification and feedback [9,30,33,36,37,39,40,42-44]
  - Personalized goals and feedback can motivate
  - Data visualization helps to plan and monitor goals
  - Health data visualization connects user to the purpose of the device
  - Poor or absent feedback diminishes value
- Emotions invoked by the device [9,36]
  - Connected to feedback
  - Relationship with the device
  - Negative feelings toward the device can lead to abandonment
  - Emotional attachment to an external motivator can be a positive driving force

- Social capital and encouragement [13,15,36,39,41,44]
  - Social capital promotes continued use
  - Wearables as adjunct to social support
  - Peer support, interaction, and communication
  - Help with troubleshooting
  - An external influence; boost motivation
- Promotion by health care staff [31,32]
  - Benefits of involvement by health sector
  - Motivated to use if part of the treatment plan
  - Input from care team to overcome barriers and meet goals

Category 3: Integration into daily life [9,13-15,34-37,39,41-43]

- Ease of integration is determined by features, day-to-day function, purpose, and reliability of the device
- Cumbersome or annoying design features hinder integration
- Lack of desired features diminishes value of the device
- The device cannot serve its function if it is unreliable and difficult to use
- Reliability issues affect routine and can lead to stigma and embarrassment
- Device issues reduce motivation and diminish the value of devices

Category 4: Device features [9,13,14,29,30,32,33,35,37,40,43]

- Preferred features (in no particular order): waterproof, step count, easy-to-read display format, GPS (security in case of a fall or getting lost), looks like a watch, comfortable location on the body (generally wrist or ankle), secure attachment, smaller, long battery life, fewer notifications, does not interfere with clothing, personalized notifications or alarms, thin and flexible band, simple attachment (easy to use with limited dexterity), comfortable to wear at night, easy to work (intelligibility), more diverse features, health-related features, tracks sleep, looks nice or cool, simple smartphone or tablet app, other activities that older adults may be doing, real-time feedback on app or device, smaller design, easy to synchronize, automatic logging of activity, goal tracking, view health information, help section, large and easy-to-press buttons, and easy to see (if falls on the floor)
- Disliked features (in no particular order): looks like a medical device (aesthetics), frequent charging, auto-goal function, inaccuracy, having to wear in bed (if uncomfortable), not capturing all activities, large and rigid band, tethered to the smartphone, uncertainties about water damage and charging, complicated tablet or smartphone, no practical training, does not match clothes, difficult to put on, frequent alarms or notifications, difficult to interact with when on the ankle, not compatible with a smartphone, difficult to handle, and not suited for older adults

The key concepts (order not indicative of prominence or salience) are (1) openness to engage and functional ability of the user; (2) motivation for device use; (3) integration into daily life; and (4) device features.

First-order quotations (raw, primary data, ie, direct participant quotations) and second-order (authors' interpretations of their primary data) interpretations were used to support the analysis of the *Translation* section. Throughout the *Results* section, first-order quotations (primary study participants) are indicated in italicized quotations and second-order interpretations (primary study authors) are indicated in italics.

Openness to Engage and Functional Ability of the User

Age-Related Physiology or Comorbidities

Certain age-related characteristics can impact users' comfort with new technologies, such as hearing loss, limited dexterity, and low vision [13,14]. Older users may be slower to process new information and therefore require simple, visible instructions:

> *Participants saw that, for senior persons less vigorous than themselves, everyday use of the device could be difficult, cumbersome, and demanding: "It is more difficult for a person less alert than me maybe also using walking aids. It might be tough for them to register like this every day." [9]*

Self-efficacy for Technology

In addition to their actual technical skills, an older person's perceptions of their technical abilities can be a barrier to adoption [13]:

> *This was reflected by many of our participants in the comments they made about the devices—often relating that they were "not built with us in mind," that they were created "for someone younger," and that devices needed a more "tech-savvy" user. [31]*

Low self-efficacy for technology can influence users' attitudes and limit the *perceived ease of use* of a device [13]:

> *"I was of course a bit worried initially about not being able to handle it. That I would push the wrong button and things like that." [9]*

Insecurities can arise when users encounter usability issues or technical failures and do not have the experience of identifying or resolving the issue:

> *The participants had felt inexperienced in handling the technical devices and therefore had felt insecure on whether they were doing this correctly. In addition, there were occasions when the technology had not worked properly, and this made the users wonder if the problems experienced were because of incorrect handling. [9]*

Individuals with higher self-efficacy for technology are more open to using wearable devices and exploring their features. Users with lower self-efficacy tend to require more support, clear instructions, and additional training to increase their sense of control and prevent device abandonment [13].

Exploration and Use of Device Features

Many older adults have a desire to learn more about their health and are interested in various advanced features (which are not always available) [13]. Sometimes, they are frustrated when their device does not have their desired features or when accessing the available features is difficult. Limited technical abilities could hinder the exploration of features and have an effect on behavior change. Clear, simple instructions help users overcome the initial technical hurdles and allow them to explore the features that they desire; this can ultimately lead to continued device use [32].

Independence

Users are less open to engaging with a device that is burdensome or limits their independence. A user will not perceive themselves as independent if they have to rely on friends, family, or researchers to help them with device issues. In addition, a device that requires frequent charging will affect the user's routine and limit the time that they can spend away from a power source [29]. Overall, if a user has to frequently seek assistance with their device, they will not be able to live an independent life with the device, which is often the goal [33]:

> *"I'm sure it's there [the support] but it means taking their time, and making my problem their problem. And that's hard for me to do because of my own attitudes about independence I think. I really resent supervision, which is intrusive and demanding; kinds of stuff like that within the family." [33]*

Motivation for Device Use

Internal Influences

For many users, a degree of motivation is required to realize behavior change or long-term use. This is not exclusive to activity trackers; some participants do not feel the need to wear a fall detection device, even if they are at risk of falling:

> *A participant who experienced four falls during the course of the trial explained he did not need the device as, "I don't consider myself a faller." [14]*

Some users felt that they were too young to need a fall detection device currently:

> *"You know if were a high fall risk...but at the moment I don't consider that. When I get old maybe." [14]*

Thus, to successfully incorporate a fall detection device into their life, the person must have a recognized need and personal desire to prioritize their safety.

Activity trackers are often worn to monitor physical activity levels. Some participants were motivated to increase their activity when the device was introduced into their lives:

> *"I was motivated by the technology, that I freely admit." [9]*

Other participants were already motivated to increase their activities before using the device:

> *"The technology has no impact on my motivation, I am physically active anyway. I am on the verge of getting diabetes, that is what motivates me the most." [9]*

Both physically active and inactive older adults can lack the motivation to use a wearable device if neither has the desire to change their activity levels. Intrinsic motivation seems to be particularly powerful for users who are inactive but have a strong desire to change this; this group has room to improve, unlike very active people who are already at their desired activity level [13]:

> *They believed that a wearable device can motivate them to improve their exercise level. This theme was more significant in seniors who did not exercise regularly and seniors with lower income. [13]*

> *Long-term users emphasized the importance of internal motivation (Just do it) where activity trackers were serving as secondary facilitators.... [37]*

Those who were already satisfied with their exercise levels saw no benefits from using the device [39]. Equally, those already motivated to exercise felt that the device had no additional effect on their behavior [13].

Quantification and Feedback

Some older adults found that quantification of their activity can drive motivation [36,37,39]. Activity tracker users often have a specific goal (eg, increase the daily step count) [36]. Devices that provide feedback help users track their progress [9,36] (eg, "I liked the ability to monitor my progress").

Each user has a different goal, so the more personalized the feedback, the more effective the device:

> *Goal setting was perceived important for increasing active behaviour: a quantitative goal was helpful for the user by clarifying if the current activity level was too low. [9]*

However, already-active individuals were not always affected by feedback:

> *"I did not change my exercise habits during the monitoring, I took the same walk as usual in the*

> *morning or in the afternoon. It is a goal I have and as a pensioner, I have plenty of time." [9]*

The feedback and features of the device must align with the goals of the user. Some users only need a push (eg, step target). Others have more detailed health monitoring goals (eg, heart rate, sleep, and quantification of multiple activities). People can feel disconnected from a device that does not provide adequate feedback; this can limit a device's ability to help the user achieve their goals.

The importance of feedback is not limited to activity trackers. Fall detection devices provide feedback in the form of alerts and calls for help. Trial participants said they would like clear feedback about when alerts were activated, who that alert was notifying, and how they could disable the alert [14,29].

Awareness of Physical Activity Levels

Using an activity tracker often leads to increased awareness of one's activity levels, particularly for those who were previously inactive:

> *"My Fitbit allowed me to personalise my exercise. I learned new things about myself from the fitbit." [13]*

However, increased awareness does not necessarily motivate the user to exercise. The desire to increase exercise levels (before knowing one's current level) and an achievable exercise goal were more motivating than awareness. Certainly, these devices can show how sedentary users are and remind them to meet their exercise goals but a person must already want to make a lifestyle change:

> *"It was more informative than motivating, because I had my own agenda that my doctor set out for me to do." [31]*

Thus, activity trackers were more often viewed as a catalyst rather than a creator of behavior change.

Emotions Invoked by Device

Feedback on activity can elicit strong emotions among users and can become attached to their results, experiencing positive affect when they meet their goals and negative affect when they do not:

> *"It was irritating when it is visible that I had been so damn lazy. But it is good to have (the technology)." [36]*

Sometimes, users are more concerned with how they connect with a device than the specific output metrics, so that emotional meaning is valued more than actual gains. When a device elicits more positive feelings than negative feelings, users are more inclined to continue use.

Devices can lead to stigma and embarrassment when drawing attention to the public:

> *"It's when they don't say anything you wonder kinda what they're looking at, cause they do take notice of it." [14]*

False alarms from fall detection devices can lead to disruptions in public [14], and activity alarms from activity trackers can interrupt meetings and social events. Aesthetically, devices with a medical look can lack acceptability, because this can draw extra attention and many older adults do not want to be viewed as a "patient":

> *"...what I was wearing was sheer, and would show this light which everybody was curious about, and it just didn't look good with, I didn't want to wear it." [14]*

Extrinsic Motivation—Social Support

For many users, the social network around the device is key to its continued use [13,36]. Social interaction and engagement and peer group attitudes toward technology are key factors influencing adoption. Ongoing peer support and encouragement can positively influence adherence:

> *"Meeting with others in the sense of: did they experience the same thing? Do they need encouragement? Can something they're doing encourage me to alter behaviours?" [36]*

Peers can also help troubleshoot issues and provide hints for maximizing user benefits [36]. Social support is important for long-term use, because intrinsic motivation can waver over time. The device can also act as its own "social support" if it provides good feedback and is easy to interact with. This is important for older adults who are isolated (physically or socially) and therefore might need help to establish a support group:

> *Long-term users indicated social support to be the main motivational factor, with the focus on building relationships around daily activity routines. Long-term users were better prepared to modify the social environment around them to maintain an active lifestyle, receive positive feedback, and seek accountability from others. [37]*

Promotion by Health Care Staff

For those using a device for medical purposes, the input and encouragement of a health professional can be important for adoption and continued use. Learning about the device from professionals can help overcome barriers to adoption and ultimately meet their goals:

> *"But if someone can guide you through it, I think any of them, once you start using them you would probably use it. But I wouldn't go to Best Buy I wouldn't have thought to go to best buy. If it's for my health, I would think to go to a pharmacy." [31]*

Integration Into Daily Life

To be successfully integrated into the user's life, the device must not only be acceptable and reliable, it must also be perceived by the user to add value to their life. The ease of integration is often determined by the purpose and features of a device and the reliability of the device's functions. Certain design features, such as appearance, weight, material, dimensions, and comfort, are particularly important. If the wearable device mimics a device already in the user's life (eg, wristwatch), it can be seamlessly integrated:

> *"Then, it becomes a habit. And this is precisely what happened to me with my watch [reference to the AT]." [44]*

If the device does not have the user's desired features (eg, swimming, activity history, or GPS tracking), the user's perceived value of the device may be low:

> *"Really, after the bloom got off the rose, I didn't like anything about it." [34]*

Conversely, users may tolerate design faults if they value the device.

A common barrier to acceptance is its unreliability. When someone cannot rely on a device to serve its purpose or give accurate feedback, it loses value, and the motivation to wear the device can wane. This is evident in the authors' conclusions:

> *Some participants...questioned whether the result was correct. This reduced their motivation for being monitored. [9]*

It is also evident in the raw participant data:

> *"...We began to think that it wasn't accurate, so it lost its appeal." [34]*

Critically, the device should not negatively affect the user's routine. Frequent charging, not being waterproof, being tethered to a smartphone, and being difficult to put on and take off are examples of features that can disrupt a user's routine [13,14]. When this happens, especially for older users, the device does not integrate into the user's life and loses its value.

Device Features

Participants across the 20 included studies generally preferred devices that have the following features: waterproof, small in size, comfortable (especially if worn at night), aesthetically pleasing (fashionable; not like a medical device), with an easy-to-read display, a long battery life, and a thin, flexible band (if worn on the wrist). They enjoyed using device features that counted their steps, tracked their location using GPS, automatically logged their activity, measured health parameters (heart rate, blood pressure, or sleep), updated them on activity goals, automatically contacted help in the event of a fall, and synchronized automatically with their other devices. They like devices that are easy to attach, are secure, do not interfere with clothing, and are easy to handle.

The participants disliked devices that were inaccurate, required frequent charging, were uncomfortable, tethered to a smartphone, were difficult to attach, were not compatible with their smartphone, were not suited for older users, or do not capture all of their daily activities. They especially disliked devices without adequate instructions to help them troubleshoot issues or turn off annoying alarms.

Synthesis

Summary of Synthesis Process

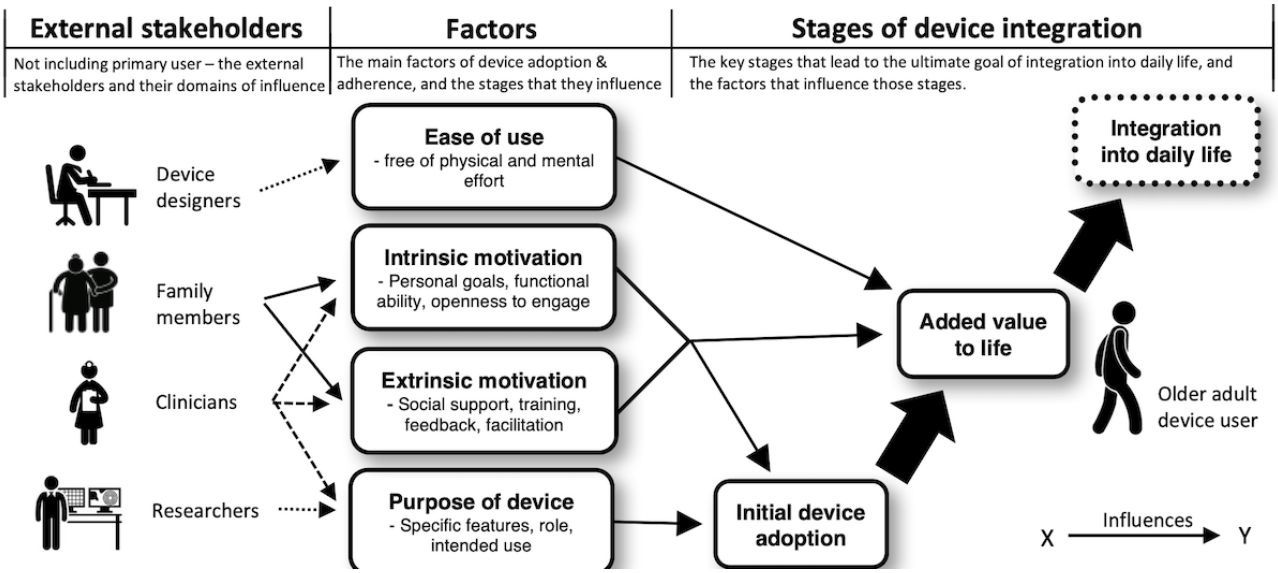
The first-order (quotations), second-order (individual themes extracted from each paper), and third-order (key concept) interpretations were considered as a whole to develop a line-of-argument synthesis.

The experience of integrating a device into everyday life is a dynamic process of assessing the added value of the device and is influenced by a range of interrelated intrinsic (internal motivation, functional ability, interest, and openness) and extrinsic (external motivation, training, device characteristics, functionality, and feedback) factors. Many factors influence whether an older adult sees a device as worth wearing, and the appraisal and balance of these factors tell the user whether the device adds value to their life.

Line of Argument

We developed a line of argument to describe the factors that influence successful integration (Figure 2). Our line of argument takes the form of a "conceptual model" of the factors that lead to the integration of the device into the user's life. Our conceptual model describes how motivation, ease of use, and device purpose determine whether a device will add value to the user's life, which subsequently determines if the device will be integrated into their life.

Figure 2. Conceptual model developed from the line-of-argument synthesis.



User motivation is key. Without motivation (eg, symptom monitoring), the user will view the device as just another piece of technology. On the basis of the data collated in this synthesis, we found that older adults do not adopt new technologies because of their novelty. We found that motivation comes in two forms: intrinsic and extrinsic motivation. They influence both the user's initial reason to adopt a device and to sustain its use.

Intrinsic motivation describes a user's personal connection to a device. Initially, the user must be motivated to make a change in their life that will be supported by adopting a wearable device (eg, increasing physical activity or detecting falls). Intrinsic motivation is often required for individuals to adopt a device initially. The device itself does not create motivation; many users commented that while being able to view their daily step count is interesting, it does not spur them to change their physical activity habits unless they are already motivated to do so. Intrinsic motivation is also important for a device to add long-term value to a user's life. It is necessary to overcome some of the usability hurdles that users face when they adopt a device. It also fuels continued use as the initial novelty wears off.

Extrinsic motivation is another important contributor to device adoption and added value. This includes factors such as training, technical support, promotion, support from health professionals, peer support, and device feedback. Initially, extrinsic motivation influences device adoption through the practicalities of acquiring and setting up the device and learning how to use its features. Older adults are often asked to adopt a device for fall detection or as part of a treatment or health regime. This form of extrinsic motivation often leads to device adoption but may not contribute to added value if other extrinsic factors (eg, technical and peer support) are not present. Technical support was frequently cited as a crucial extrinsic motivator, both initially and over time. Good technical support connotes added value because it gives the user the confidence to explore the device's features and supports integration into the user's life. Peer support is another important extrinsic motivation factor that contributes to both device adoption and added value. It comes in many forms and is unique for each user, but its importance is universal. Social support encourages older adults to adopt wearable devices and motivates continued use. Conversely, reliance on social support (eg, having to bother someone for assistance) limits the user's independence and could be a barrier to continued use. Together, the factors of intrinsic and extrinsic motivation influence whether an individual will adopt a device and whether the device will continue to add value to their life.

The purpose of the device (fall detection, step count, etc) is the main reason why older adults adopt it, and it is what initially draws a user to that specific device. Unlike those of a younger generation, older adults do not tend to use new technology simply because they have fun features. They view devices as tools and expect them to serve their purposes with accuracy and reliability. The purpose of a device (and its features) is key to adoption. Older adults are unlikely to adopt a device that does not fulfill pre-existing needs. The purpose of the device also adds value and facilitates integration as the user expands their relationship with the device. Upon adoption, the user evaluates whether, and to what extent, the device serves its intended purpose. A device that continues to serve its intended purpose (or serves additional purposes as the user becomes more familiar with its features) is perceived to add value, leading to integration as the user relies on that device to fulfill an important need in their life.

Along with motivation and purpose, ease of use also predicts added value. This is defined as the degree to which a device is free of physical and mental effort for users. Specific device features (eg, battery life and touch screen menus) influence ease of use, as do general features such as access to simple instructions and the amount of interaction required. An easy-to-use device adds value by reducing the burden of using the device. This allows users to focus on their motivators and the fundamental purpose of the device.

Added value to life is the ultimate contributor to successful integration into daily life. The added value is the resulting balance of motivators (or lack thereof), device features (and their accuracy), ease of use, device purpose, and user experience. When the negatives outweigh the positives, the device will most likely not be integrated into everyday life.

Discussion

Principal Findings

This is the first study to systematically review and synthesize the qualitative literature on older adults' experiences with wearable devices. This meta-synthesis collated the experiences of 349 trial participants and presented the key factors that influence user acceptance and adherence. These factors include intrinsic and extrinsic motivation to use the device; the purpose of the device and how it relates to the user's expectations and needs; and the ease of use and functional ability of the user. The user's appraisal of these factors determines the level of value added by the device to the life of each user.

Motivation for device use comes in two forms (intrinsic and extrinsic) and encompasses many aspects of the user experience. According to our line-of-argument synthesis, motivation influences both device adoption and added value. Motivation seems to be as, if not more, important for older adults than the actual device features. Moreover, the user's needs and the support structure around the device—aspects that are often overlooked—seem to play a crucial role in long-term adoption.

Comparisons With Previous Work

Our bottom-up inductive qualitative synthesis supports the findings of existing theory-bound models of technology acceptance. It was not intended from the outset that our conceptual model would tie in with quantitative models such as the Technology Acceptance Model (TAM) and the value-based adoption model (VAM). We felt that the TAM and VAM could be used to structure and contextualize our findings. These models use quantitative surveys to test hypotheses about factors that predict the intention to use. For example, the TAM shows that perceived ease of use, perceived usefulness, and attitude toward the system predict intention to use [51]. The development of these models does not involve trial components or qualitative methods. Therefore, our study should not be

directly compared with the TAM or any other acceptance model and should instead provide inspiration for hypotheses to test in future iterations of wearable device acceptance models.

Originally designed to describe the acceptance of information services in organizations [51], the TAM has only recently been applied to wearable devices [52-55]. In contrast to the TAM, the authors of the VAM recognized that most consumers adopt mobile technologies for personal purposes and that the cost of voluntary adoption is borne by the individual, not the organization [56]. According to the VAM, perceived sacrifices (cost and technicality) seem to have a greater impact than perceived benefits (usefulness and enjoyment) on perceived value. Many participants in our review alluded to this balance between sacrifices and benefits. They described how a device that disrupts their routine or limits their independence is not worth the hassle, especially if the features (eg, counting steps) are not beneficial to them.

The TAM and VAM inspired the recent development of a smart wearables acceptance model for older adults by Li et al [57]. Along with established acceptance factors such as perceived usefulness [51], Li et al [57] included older adult–specific factors such as self-reported health conditions, perceived social risk, performance risk, and social influence. Their results supported their hypothesis that facilitating conditions positively predicted intention to use, a finding supported by our study. They also showed that the self-reported health status is a negative predictor of use, suggesting that older adults with a better health status are not likely to require these technologies. Our results show a similar trend; older adults are motivated to use a device if they have a need (eg, monitor symptoms and fall risk). In contrast to our findings, 95.9% (140/146) of their participants perceived minimal or no social risk when using a wearable device. This may be because their participants were not offered the opportunity to wear the device in public. Several of the participants in our review (who did wear their devices in public) described feelings of embarrassment or stigma when the device intrudes on the user's life.

Our review points to age-related factors that can influence acceptance, such as experience with technology and openness to engage. A systematic review of factors influencing acceptance of technology for ageing-in-place found a similar phenomenon and discussed the effect of age and chronic illness on the acceptance of vital sign monitoring systems [58,59]. The systematic review also highlighted the impact of social support from family, friends, professional caregivers, and peers [58]. In their 2014 review of determinants and barriers, Lee and Coughlin [60] described eight similar factors (value, usability, technical support, social support, emotions, independence, experience, and confidence) and two additional factors that our review did not uncover (affordability and accessibility) [60]. Similar findings from these studies indicate a convergence of the field toward an understanding of the key factors that influence adoption and adherence.

Relevance for Researchers, Clinicians, and Designers

When designing future wearable device acceptance models for older adults, researchers should consider the multiple stages of device use that follow the initial "intention to use." Furthermore, in the user experience, concepts such as added value become relevant, which may have a different set of predictors than the initial intention to use. In the development of the Senior Technology Acceptance and Adoption Model (STAM), Renaud and van Biljon [61] related certain acceptance factors to adoption stages. The STAM includes factors such as *confirmed usefulness* and *ease of learning and use*. Using qualitative methods to develop their model, Renaud and van Biljon [61] used these factors to explain why older adults do not reach the final adoption phase and never fully accept technology. Yu-Huei [62] adapted the STAM to wearable devices and added two additional factors, information source and group behavior, which emerged from their qualitative analysis of older adults in Taiwan.

Our study proposed several predictors that could be tested as a part of future model development studies. First, motivation is key and seems to be a constant driving force throughout the user experience. Researchers should question users on both intrinsic and extrinsic motivators to see if these factors predict integration. Second, the purpose of the device (and whether the user aligns with that purpose) should be investigated as a predictor of device adoption and added value. Finally, ease of use should be considered within the context of older adults, as done by Yu-Huei [62]. The technical abilities of older adults may differ from those of younger generations, and certain medical conditions may hinder functional abilities.

Researchers play a role in validating (or refuting) the findings of this review. Older adults require a specific wearable device acceptance model because they are a distinct population from the individuals used to develop the TAM and other wearable device acceptance models [54]. As opposed to smartphones and embodied conversational agents, wearable devices are designed to be used individually and continuously, which leads to a unique set of influencing factors.

For clinical trial researchers, we stress the need to provide extrinsic motivation for their participants by conveying the importance of the device. They should also provide training and technical support to facilitate ease of use. Clinicians using wearable devices in their private practice can provide extrinsic motivation by clearly explaining the device's purpose and the meaning of measurements to their patients. They should also provide encouragement and technical support. They can support intrinsic motivation by learning about their patients' health goals and desires to use a wearable device.

Although not explicitly stated, several of the included studies used a user-centered design approach [9,14,29]. A user-centered design, though often overlooked, is essential in the development of wearable devices, particularly if the designers are not a part of the intended demographic. Designers, clinicians, and researchers can collaborate with older adults to address all aspects of wearable device use. As the studies in our review demonstrate, trials and qualitative data are valuable tools for designers using the user-centered design approach to explore the long-term use of their products. Some device features become relevant only after a period of usage. For the benefit of designers, we summarized some of the features that were commonly discussed in our review. The study participants discussed many design features that are specific to the devices

in question, and thus features could not be compared between the studies. However, we can make a few generalizations. First, designers should not take for granted that older adults will accept every design feature. Second, a device should serve the purpose, and the primary function of the device should be reliable and easy to use; some older adults are interested in advanced features, but not all are. Third, participants disliked unnecessary interactions with the device, such as frequent charging, responding to alerts, and maintaining a Bluetooth connection to their smartphone. An ideal device requires little maintenance and only requires interaction to monitor the data and obtain device feedback. Finally, a wearable device should be easy to take on and off, comfortable to wear at night, and waterproof. It should cause minimal disruption in users' lives, be aesthetically pleasing, and should not draw attention to the user or single them out as a patient; it should not interfere with clothing; and it should have a silent mode to prevent unnecessary disruptions.

Strengths, Limitations, and Future Directions

Measures such as practicing reflexivity and using 2 reviewers maximized the quality of this meta-synthesis. The authors are an experienced multidisciplinary team (geriatric medicine, psychology, epidemiology, and engineering) with expertise in qualitative approaches. By following the eMERGe reporting guidance (see Table S1 in Multimedia Appendix 2 [50]), we communicated our methodology with transparency, including our study's limitations. Although both reviewers (KM and LK) collaborated to generate the search strategy and inclusion criteria, because the second reviewer only screened a sample of the references, some relevant studies may have been excluded during the screening process. Our database search only included 70% (14/20) of the included studies. The additional 6 studies were located by manually searching the reference lists and searching Google Scholar, highlighting the limitations of our search strategy. For example, we should have included the search term "senior*" to find studies such as the Abouzahra and Ghasemaghaei [13] study. It is worth noting that this issue is more common in qualitative reviews than in quantitative reviews because of the poor and inconsistent indexing of qualitative research in databases [63]. While we searched four databases, searching for additional databases (eg, Scopus and ISI Web of Science) would have strengthened our study.

The content of our results and the line-of-argument conceptual model were contingent on the data collected from the broad inclusion criteria. This is both a strength and a weakness of qualitative syntheses; it affords reviewers the flexibility to uncover new ideas but it dictates which questions can be answered. Some studies lacked rich descriptions and interpretations of their findings or did not provide sufficient context (about the sample, the device, or the procedures). This limited the contributions of some studies to the meta-synthesis [64], regardless of their ETQS quality score. Studies with low ETQS scores often fell short of simple aspects, such as not reporting the location of the study. Ultimately, the results of this study are based on the synthesis of qualitative data, which is inherently subjective. Our line-of-argument and conceptual model provide suggestions, but the full development of a wearable device acceptance model for older adults will take place with a more rigorous study design.

All studies included in this review were published in English and conducted in Western countries. The findings may not represent countries with different cultures, access to wearables, and income levels. Several studies included short trial periods and a few participants. Each study evaluated a different device (or set of devices), which restricted comparisons between studies. Future research would benefit from long-term trials using in-depth qualitative methods to evaluate the drivers of acceptance and adherence. Future research must also include the views of older adults who use wearable devices as part of clinical care, not just a research trial. An alternative set of predictors might be relevant to participants who use a device for a specific health purpose.

Conclusions

This review found that several key factors influence the acceptance and use of wearable devices by older adults. These include intrinsic and extrinsic motivation for device use, ease of use, device purpose, and perceived added value to the user's life. Designers, clinicians, and researchers should be aware that useful device features alone do not lead to continued use. To overcome the usability barriers (eg, limited technical ability), an older adult must be motivated to use a device because it serves a useful purpose. A support structure should be placed around the user that fosters motivation, encourages engagement with peers, and adapts to the user's preferences. Future research should evaluate our conceptual model by validating our proposed predictors and conducting long-term wearable device trials that use qualitative methods to comprehensively address the multiple stages of device use and the many factors that contribute to adherence.

Acknowledgments

This review was cofunded by the European Regional Development Fund under Ireland's European Structural and Investment Fund Programme 2014-2020. This work was supported in part by the Interreg Norther Periphery and Artic Programme funded project SENDoc (Smart sENsor Devices fOr rehabilitation and Connected health). Aspects of this publication were supported by Enterprise Ireland and Abbvie Inc under grant IP 2017 0625.

Authors' Contributions

KM, EO, and S Timmons conceived the study. KM and EO guided the study methodology and execution. KM and LK collected and analyzed the data. KM drafted the manuscript. All authors (KM, EO, LK, JB, S Tedesco, MS, CC, AA, JC, AN, and S Timmons) critically reviewed and provided intellectual input to the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation tool for the qualitative studies quality appraisal checklist.
[DOCX File , 96 KB - mhealth_v9i6e23832_app1.docx]

Multimedia Appendix 2
The Meta-Ethnography Reporting Guidance checklist.
[DOCX File , 116 KB - mhealth_v9i6e23832_app2.docx]

References

1. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. J Am Med Assoc 2016 Aug 02;316(5):538-540. [doi: 10.1001/jama.2016.9124] [Medline: 27483069]
2. Active Ageing: A Policy Framework. World Health Organization. 2002. URL: https://apps.who.int/iris/bitstream/handle/10665/67215/WHO_NMH_NPH_02.8.pdf;jsessionid=05C9853E78DCB8F5182322FAF23C4798?sequence=1 [accessed 2021-05-11]
3. Malwade S, Abdul SS, Uddin M, Nursetyo AA, Fernandez-Luque L, Zhu X, et al. Mobile and wearable technologies in healthcare for the ageing population. Comput Methods Programs Biomed 2018 Jul;161:233-237. [doi: 10.1016/j.cmpb.2018.04.026] [Medline: 29852964]
4. Baig MM, GholamHosseini H, Moqeem AA, Mirza F, Lindén M. A systematic review of wearable patient monitoring systems - current challenges and opportunities for clinical adoption. J Med Syst 2017 Jul;41(7):115. [doi: 10.1007/s10916-017-0760-1] [Medline: 28631139]
5. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
6. Sun N, Rau PP. The acceptance of personal health devices among patients with chronic conditions. Int J Med Inform 2015 Apr;84(4):288-297. [doi: 10.1016/j.ijmedinf.2015.01.002] [Medline: 25655783]
7. Nehra AK, Gettman MT, Rivera ME, Agarwal DK, O'Neil DA, Jenkins SM, et al. A survey of perceptions and acceptance of wearable technology for health monitoring in a urological patient population. Urol Pract 2017 Nov;4(6):508-514. [doi: 10.1016/j.urpr.2016.10.002]
8. Fisher JM, Hammerla NY, Ploetz T, Andras P, Rochester L, Walker RW. Unsupervised home monitoring of Parkinson's disease motor symptoms using body-worn accelerometers. Parkinsonism Relat Disord 2016 Dec;33:44-50. [doi: 10.1016/j.parkreldis.2016.09.009] [Medline: 27637282]
9. Ehn M, Eriksson LC, Åkerberg N, Johansson A. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]
10. Steele R, Lo A, Secombe C, Wong YK. Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. Int J Med Inform 2009 Dec;78(12):788-801. [doi: 10.1016/j.ijmedinf.2009.08.001] [Medline: 19717335]
11. Ozanne A, Johansson D, Graneheim UH, Malmgren K, Bergquist F, Alt Murphy M. Wearables in epilepsy and Parkinson's disease - a focus group study. Acta Neurol Scand 2018 Feb;137(2):188-194. [doi: 10.1111/ane.12798] [Medline: 28714112]
12. Steinert A, Haesner M, Steinhagen-Thiessen E. Activity-tracking devices for older adults: comparison and preferences. Univ Access Inf Soc 2017 Apr 8;17(2):411-419. [doi: 10.1007/s10209-017-0539-7]
13. Abouzahra M, Ghasemaghaei M. The antecedents and results of seniors' use of activity tracking wearable devices. Health Policy Technol 2020 Jun;9(2):213-217 [FREE Full text] [doi: 10.1016/j.hlpt.2019.11.002]
14. Demiris G, Chaudhuri S, Thompson HJ. Older adults' experience with a novel fall detection device. Telemed J E Health 2016 Sep;22(9):726-732 [FREE Full text] [doi: 10.1089/tmj.2015.0218] [Medline: 26959299]
15. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriatr Psychiatry 2019 Apr;34(4):617-624. [doi: 10.1002/gps.5064] [Medline: 30701592]
16. Atkins S, Lewin S, Smith H, Engel M, Fretheim A, Volmink J. Conducting a meta-ethnography of qualitative literature: lessons learnt. BMC Med Res Methodol 2008 Apr 16;8:21 [FREE Full text] [doi: 10.1186/1471-2288-8-21] [Medline: 18416812]

17. Noblit GW, Hare RD. Meta-Ethnography: Synthesizing Qualitative Studies. Thousand Oaks, California, United States: Sage Publications Inc; 1988.
18. Kekade S, Hseieh C, Islam MM, Atique S, Khalfan AM, Li Y, et al. The usefulness and actual use of wearable devices among the elderly population. Comput Methods Programs Biomed 2018 Jan;153:137-159. [doi: 10.1016/j.cmpb.2017.10.008] [Medline: 29157447]
19. Kristoffersson A, Lindén M. A systematic review on the use of wearable body sensors for health monitoring: a qualitative synthesis. Sensors (Basel) 2020 Mar 09;20(5):1502 [FREE Full text] [doi: 10.3390/s20051502] [Medline: 32182907]
20. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and iot-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jun 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
21. Shin G, Jarrahi MH, Fei Y, Karami A, Gafinowitz N, Byun A, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: a systematic literature review. J Biomed Inform 2019 May;93:103153. [doi: 10.1016/j.jbi.2019.103153] [Medline: 30910623]
22. Son H, Park WS, Kim H. Mobility monitoring using smart technologies for Parkinson's disease in free-living environment. Collegian 2018 Oct;25(5):549-560. [doi: 10.1016/j.colegn.2017.11.005]
23. Thordardottir B, Fänge AM, Lethin C, Gatta DR, Chiatti C. Acceptance and use of innovative assistive technologies among people with cognitive impairment and their caregivers: a systematic review. Biomed Res Int 2019;2019:1-18 [FREE Full text] [doi: 10.1155/2019/9196729] [Medline: 30956989]
24. Cooper C, Gross A, Brinkman C, Pope R, Allen K, Hastings S, et al. The impact of wearable motion sensing technology on physical activity in older adults. Exp Gerontol 2018 Oct 02;112:9-19 [FREE Full text] [doi: 10.1016/j.exger.2018.08.002] [Medline: 30103025]
25. Allet L, Knols RH, Shirato K, de Bruin ED. Wearable systems for monitoring mobility-related activities in chronic disease: a systematic review. Sensors (Basel) 2010;10(10):9026-9052 [FREE Full text] [doi: 10.3390/s101009026] [Medline: 22163393]
26. Mahoney EL, Mahoney DF. Acceptance of wearable technology by people with Alzheimer's disease: issues and accommodations. Am J Alzheimers Dis Other Demen 2010 Sep;25(6):527-531 [FREE Full text] [doi: 10.1177/1533317510376944] [Medline: 20702501]
27. Zimmer L. Qualitative meta-synthesis: a question of dialoguing with texts. J Adv Nurs 2006 Feb;53(3):311-318. [doi: 10.1111/j.1365-2648.2006.03721.x] [Medline: 16441536]
28. Bergmann JH, McGregor AH. Body-worn sensor design: what do patients and clinicians want? Ann Biomed Eng 2011 Sep;39(9):2299-2312. [doi: 10.1007/s10439-011-0339-9] [Medline: 21674260]
29. Thilo FJ, Hahn S, Halfens RJ, Schols JM. Usability of a wearable fall detection prototype from the perspective of older people - a real field testing approach. J Clin Nurs 2019 Jan;28(1-2):310-320. [doi: 10.1111/jocn.14599] [Medline: 29964344]
30. Floegel TA, Allen KD, Buman MP. A pilot study examining activity monitor use in older adults with heart failure during and after hospitalization. Geriatr Nurs 2019;40(2):185-189. [doi: 10.1016/j.gerinurse.2018.10.001] [Medline: 30391041]
31. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
32. Nguyen NH, Hadgraft NT, Moore MM, Rosenberg DE, Lynch C, Reeves MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. Support Care Cancer 2017 Nov;25(11):3375-3384. [doi: 10.1007/s00520-017-3756-y] [Medline: 28540402]
33. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
34. Fausset CB, Mitzner TL, Price CE, Jones BD, Fain BW, Rogers WA. Older adults' use of and attitudes toward activity monitoring technologies. Proc Hum Factors Ergon Soc Annu Meet 2013 Sep;57(1):1683-1687 [FREE Full text] [doi: 10.1177/1541931213571374] [Medline: 31263349]
35. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older adults' acceptance of activity trackers. J Appl Gerontol 2017 Feb;36(2):127-155 [FREE Full text] [doi: 10.1177/0733464815624151] [Medline: 26753803]
36. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a wearable activity device in rural older obese adults: a pilot study. Gerontol Geriatr Med 2016;2:1-6 [FREE Full text] [doi: 10.1177/2333721416678076] [Medline: 28138502]
37. Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
38. Hermanns M, Haas BK, Lisk J. Engaging older adults with Parkinson's disease in physical activity using technology: a feasibility study. Gerontol Geriatr Med 2019;5:2333721419842671 [FREE Full text] [doi: 10.1177/2333721419842671] [Medline: 31069250]

39. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: 10.2196/12346] [Medline: 31625951]
40. Lee BC, Ajisafe TD, Van Thanh Vo T, Xie J. Understanding long-term adoption and usability of wearable activity trackers among active older adults. In: Human Aspects of IT for the Aged Population. Design for the Elderly and Technology Acceptance. Switzerland: Springer; 2019:238-249.
41. Rosales A, Fernández-Ardèvol M, Ferran-Ferrer N. Long-term appropriation of smartwatches among a group of older people. In: Human Aspects of IT for the Aged Population. Acceptance, Communication and Participation. Switzerland: Springer; 2018:135-148.
42. Schlomann A. A case study on older adults' long-term use of an activity tracker. Gerontechnology 2017 Jun 30;16(2):115-124. [doi: 10.4017/gt.2017.16.2.007.00]
43. Schlomann A, von Storch K, Rasche P, Rietz C. Means of motivation or of stress? The use of fitness trackers for self-monitoring by older adults. HBScience 2016 Jun 13;7(3):111-116. [doi: 10.1007/s16024-016-0275-6]
44. Zhou S, Ogihara A, Nishimura S, Jin Q. Analyzing the changes of health condition and social capital of elderly people using wearable devices. Health Inf Sci Syst 2018 Dec;6(1):4 [FREE Full text] [doi: 10.1007/s13755-018-0044-2] [Medline: 29692887]
45. Long AF, Godfrey M. An evaluation tool to assess the quality of qualitative research studies. Int J Soc Res Methodol 2004 Apr;7(2):181-196. [doi: 10.1080/1364557032000045302]
46. Hannes K, Lockwood C, Pearson A. A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. Qual Health Res 2010 Dec;20(12):1736-1743. [doi: 10.1177/1049732310378656] [Medline: 20671302]
47. CASP qualitative checklist 2018. Critical Appraisal Skills Programme. URL: https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf [accessed 2021-05-11]
48. France EF, Ring N, Thomas R, Noyes J, Maxwell M, Jepson R. A methodological systematic review of what's wrong with meta-ethnography reporting. BMC Med Res Methodol 2014 Nov 19;14:119 [FREE Full text] [doi: 10.1186/1471-2288-14-119] [Medline: 25407140]
49. Toye F, Seers K, Allcock N, Briggs M, Carr E, Barker K. Meta-ethnography 25 years on: challenges and insights for synthesising a large number of qualitative studies. BMC Med Res Methodol 2014 Jun 21;14:80 [FREE Full text] [doi: 10.1186/1471-2288-14-80] [Medline: 24951054]
50. France EF, Cunningham M, Ring N, Uny I, Duncan EA, Jepson RG, et al. Improving reporting of meta-ethnography: the eMERGe reporting guidance. BMC Med Res Methodol 2019 Jan 31;19(1):25 [FREE Full text] [doi: 10.1186/s12874-018-0600-0] [Medline: 30709371]
51. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man Mach Stud 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
52. Lunney A, Cunningham NR, Eastin MS. Wearable fitness technology: a structural investigation into acceptance and perceived fitness outcomes. Comp Hum Beh 2016 Dec;65:114-120. [doi: 10.1016/j.chb.2016.08.007] [Medline: 30872122]
53. Chuah SH, Rauschnabel PA, Krey N, Nguyen B, Ramayah T, Lade S. Wearable technologies: the role of usefulness and visibility in smartwatch adoption. Comp Hum Beh 2016 Dec;65:276-284. [doi: 10.1016/j.chb.2016.07.047]
54. Kim KJ, Shin D. An acceptance model for smart watches: implications for the adoption of future wearable technology. Internet Res 2015;25(4):527-541. [doi: 10.1108/IntR-05-2014-0126]
55. Chang HS, Lee SC, Ji YG. Wearable device adoption model with TAM and TTF. Int J Mob Commun 2016;14(5):518. [doi: 10.1504/ijmc.2016.078726]
56. Kim H, Chan HC, Gupta S. Value-based adoption of mobile internet: an empirical investigation. Decis Support Syst 2007 Feb;43(1):111-126. [doi: 10.1016/j.dss.2005.05.009]
57. Li J, Ma Q, Chan AH, Man S. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
58. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
59. Lai CK, Chung JC, Leung NK, Wong JC, Mak DP. A survey of older Hong Kong people's perceptions of telecommunication technologies and telecare devices. J Telemed Telecare 2010;16(8):441-446. [doi: 10.1258/jtt.2010.090905] [Medline: 20870685]
60. Lee C, Coughlin JF. Perspective: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
61. Renaud K, van Biljon J. Predicting technology acceptance and adoption by the elderly: a qualitative study. In: Proceedings of the 2008 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on It Research in Developing Countries: Riding the Wave of Technology. 2008 Presented at: SAICSIT '08: 2008 Annual Conference of the South African Institute of Computer Scientists and Information Technologists; October, 2008; Wilderness South Africa p. 210-219. [doi: 10.1145/1456659.1456684]

62. Yu-Huei C, Ja-Shen C, Ming-Chao W. Why do older adults use wearable devices: a case study adopting the Senior Technology Acceptance Model (STAM). In: Proceedings of the Portland International Conference on Management of Engineering and Technology (PICMET). 2019 Presented at: Portland International Conference on Management of Engineering and Technology (PICMET); Aug. 25-29, 2019; Portland, OR, USA. [doi: 10.23919/picmet.2019.8893767]
63. Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. Syst Rev 2016 May 04;5:74 [FREE Full text] [doi: 10.1186/s13643-016-0249-x] [Medline: 27145932]
64. Walsh D, Downe S. Meta-synthesis method for qualitative research: a literature review. J Adv Nurs 2005 Apr;50(2):204-211. [doi: 10.1111/j.1365-2648.2005.03380.x] [Medline: 15788085]

Abbreviations

ETQS: Evaluation Tool for Qualitative Studies
SENDoc: Smart Sensor Devices for Rehabilitation and Connected Healthy
eMERGe: Meta-Ethnography Reporting Guidance
STAM: Senior Technology Acceptance and Adoption Model
TAM: Technology Acceptance Model
VAM: value-based adoption model

Edited by L Buis; submitted 24.09.20; peer-reviewed by D Wu, V Stara, M Behzadifar; comments to author 18.10.20; revised version received 19.01.21; accepted 26.03.21; published 03.06.21.

Please cite as:
Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, Crowe C, Alamäki A, Condell J, Nordström A, Timmons S
Older Adults' Experiences With Using Wearable Devices: Qualitative Systematic Review and Meta-synthesis
JMIR Mhealth Uhealth 2021;9(6):e23832
URL: https://mhealth.jmir.org/2021/6/e23832
doi:10.2196/23832
PMID:34081020

Review

Mobile and Wearable Technology for the Monitoring of Diabetes-Related Parameters: Systematic Review

Ciro Rodriguez-León[1,2], BSc, MSc; Claudia Villalonga[1], BSc, MSc, PhD; Manuel Munoz-Torres[3,4,5], PhD, MD; Jonatan R Ruiz[6], BSc, MSc, PhD; Oresti Banos[1], BSc, MSc, PhD

[1]Research Center for Information and Communication Technologies, University of Granada, Granada, Spain

[2]Department of Computer Science, University of Cienfuegos, Cienfuegos, Cuba

[3]Departament of Medicine, University of Granada, Granada, Spain

[4]Endocrinology and Nutrition Unit, Hospital Universitario Clinico San Cecilio, Granada, Spain

[5]Centro de Investigación Biomédica en Red sobre Fragilidad y Envejecimiento Saludable, Instituto de Salud Carlos III, Madrid, Spain

[6]PROmoting FITness and Health through Physical Activity Research Group, Department of Physical Education and Sports, University of Granada, Granada, Spain

Corresponding Author:
Oresti Banos, BSc, MSc, PhD
Research Center for Information and Communication Technologies
University of Granada
C/ Periodista Rafael Gomez, 2
Granada, 18071
Spain
Phone: 34 958248598
Fax: 34 958248993
Email: oresti@ugr.es

Abstract

Background: Diabetes mellitus is a metabolic disorder that affects hundreds of millions of people worldwide and causes several million deaths every year. Such a dramatic scenario puts some pressure on administrations, care services, and the scientific community to seek novel solutions that may help control and deal effectively with this condition and its consequences.

Objective: This study aims to review the literature on the use of modern mobile and wearable technology for monitoring parameters that condition the development or evolution of diabetes mellitus.

Methods: A systematic review of articles published between January 2010 and July 2020 was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Manuscripts were identified through searching the databases Web of Science, Scopus, and PubMed as well as through hand searching. Manuscripts were included if they involved the measurement of diabetes-related parameters such as blood glucose level, performed physical activity, or feet condition via wearable or mobile devices. The quality of the included studies was assessed using the Newcastle-Ottawa Scale.

Results: The search yielded 1981 articles. A total of 26 publications met the eligibility criteria and were included in the review. Studies predominantly used wearable devices to monitor diabetes-related parameters. The accelerometer was by far the most used sensor, followed by the glucose monitor and heart rate monitor. Most studies applied some type of processing to the collected data, mainly consisting of statistical analysis or machine learning for activity recognition, finding associations among health outcomes, and diagnosing conditions related to diabetes. Few studies have focused on type 2 diabetes, even when this is the most prevalent type and the only preventable one. None of the studies focused on common diabetes complications. Clinical trials were fairly limited or nonexistent in most of the studies, with a common lack of detail about cohorts and case selection, comparability, and outcomes. Explicit endorsement by ethics committees or review boards was missing in most studies. Privacy or security issues were seldom addressed, and even if they were addressed, they were addressed at a rather insufficient level.

Conclusions: The use of mobile and wearable devices for the monitoring of diabetes-related parameters shows early promise. Its development can benefit patients with diabetes, health care professionals, and researchers. However, this field is still in its early stages. Future work must pay special attention to *privacy and security* issues, the use of new emerging sensor technologies, the combination of mobile and clinical data, and the development of validated clinical trials.

(JMIR Mhealth Uhealth 2021;9(6):e25138) doi:10.2196/25138

KEYWORDS

diabetes; monitoring; passive sensing; smartphone; wearable; mobile phone

Introduction

Background

Diabetes mellitus (DM) is a metabolic disorder primarily characterized by high blood glucose levels (GLs). People with DM are more likely to have other major health problems. Therefore, the chances for them to require special medical attention increases as the patients' quality of life decreases [1]. Statistics from the International Diabetes Federation show that 463 million people had DM worldwide in 2019, with a total of 4.2 million estimated deaths that year. The projection of these data is alarming for the next years, and by 2045, an increase of 33.9% is estimated; thus, 700 million people would have DM [2]. Such a continued increase in the prevalence of DM is mainly justified by the global rise in obesity, driven foremost by people's unwholesome lifestyles and urbanization [3]. According to the International Diabetes Federation, more people have died from DM than from other diseases sometimes categorized as more dangerous or receiving more attention from health agencies or governments [4].

DM is normally categorized into 3 groups: type 1 diabetes (T1D), type 2 diabetes (T2D), and gestational diabetes (GD). T1D affects between 5% and 10% of patients with DM and most often occurs in young people [5,6]. T1D is fundamentally characterized by a severe problem of insulin secretion. Patients are required to use an external source of insulin to balance their blood GLs. These multiple daily doses can be administered through injections or continuous insulin pumps. T2D affects between 90% and 95% of patients with DM, usually adults and senior citizens [5,6]. In this case, the pathophysiological mechanism is insulin resistance, and over time, the body loses the ability to secrete the right amount of this hormone. These type of patients with DM have several options to treat their condition.

Irrespective of the type of DM, low or too high GLs in the blood for long periods can induce several complications in patients, leading to premature death in worst cases. Critical hypoglycemia can cause comatose states and induce seizures. Chronic hyperglycemia can cause vascular damage; affect the heart, kidneys, eyes, and nerves; and lead to other serious complications [7,8]. These complications of DM can be classified into 2 types: microvascular (related to retinopathies, nephropathies, and neuropathies) and macrovascular (mainly related to cardiovascular problems) [9,10].

Extensive tests have proven that appropriate metabolic control in all DM types can delay the onset and evolution of its complications [10]. In addition, early diagnosis, continuous health care, and adequate self-monitoring of the disease by patients are key for preventing or minimizing complications. Moreover, several studies indicate that T2D can be prevented by maintaining a healthy lifestyle, attaining adequate nutrition, performing physical activity, and avoiding obesity [11]. Therefore, some parameters are of special interest to be monitored by patients with DM, such as body weight, GL, performed physical activity, blood pressure, low-density lipoprotein cholesterol, triglycerides, microalbuminuria, glycated hemoglobin level [12], acquired calories, feet condition, eye conditions, and stress levels, among others [4,13]. For all issues related to DM, governments and institutions must urgently implement new strategies, primarily fostered by the potential of diabetes technologies, to decrease the risk factors that lead to T2D and guarantee quality health care for people with DM [14].

The health community uses the term *diabetes technology* for devices and software that patients with DM use to aid their condition. According to this classification, diabetes technology has 2 main categories: insulin administration and blood glucose monitoring. Insulin pumps are the most popular devices used for insulin administration. Continuous glucose monitors (CGMs) are most often used for monitoring blood GLs. In recent years, hybrid devices have been developed, including both functions. When this technology is used properly, it can improve the quality of life of patients. However, the complexity and rapid changes in this field can be an obstacle to their widespread use by patients [15]. Although diabetes technologies have been mostly dominated by this type of devices, it is only recently that the advances in mobile and wearable technologies have burst in to complement prior diabetes technology with a new generation of digital solutions at the reach of most people.

The widespread adoption of wearable and mobile technologies around the world offers new opportunities for researchers to provide medical care and information in a portable and affordable way [16]. Smartphones stand out because of their strong computational features and pervasiveness. As of 2019, 2.5 billion people owned smartphones [17], which is far beyond the number of desktops or laptop computer ownership in countries such as the United States [18]. Even among older adults, smartphone ownership has doubled. By the beginning of 2019, there were approximately 350.4 million people using wearable technology [19], namely, devices worn directly on or loosely attached to a person [20], with a growing trend in the use of wrist-worn devices [21]. This technology has distinctive applications in the health care field because of its capacity to gather, store, and transmit data and sometimes even process it. Both patients and physicians can leverage these features for the management, treatment, and assessment of their conditions. Despite the aforementioned capabilities of mobile and wearable technology, these devices are not yet extensively used in clinical settings [20].

Several sensors are readily available on regular smartphones, such as the accelerometer (ACC), GPS, camera, ambient light, and microphone, among others. The data collected by these sensors can be used to determine the user context [22]. For example, physical activity or calories burned by the user can normally be measured using a smartphone's motion sensors

(ACC and gyroscope) [23]. Wearables have features that increasingly compare to those of regular smartphones, including, in some cases, built-in GPS, barometer, heart rate (HR), ACC, or gyroscope sensors. In addition, wearables outmatch smartphones while sensing physiological signs, such as HR, electrocardiogram (ECG), or skin temperature, which are considered of particular interest for the monitoring of DM-related parameters. Some of these physiological measurement capabilities can also be instrumented in smartphones via external pluggable devices, although only for occasional use [24]. By processing the collected data generated by the sensors on smartphones and wearables, it is possible to monitor many of the relevant parameters for patients with DM, such as GL, blood pressure, calories, physical activity, feet condition, eye condition, and stress levels. In addition, one of the most relevant characteristics of this technology is its capacity to monitor in a continuous, passive, and unobtrusive way, without necessarily interfering with people's regular daily living.

Mobile and wearable devices generate an enormous amount of data, and their ability to process these data is beyond human skills [25]. This is why sophisticated mechanisms such as artificial intelligence (AI) are most often used in combination with these devices to digest and extract meaningful knowledge from the gathered data. AI is widely used to support advanced analytics and provide individualized medical assistance [26]. In addition, a growing number of health care companies are applying AI algorithms to discover relevant clinical information from large amounts of data [27]. The main reasons for this growth include the explosive increase in the amount of data available, along with the improved performance of intelligent methodologies capable of handling and processing it. AI is also attracting great attention to DM, as the amount of data acquired electronically by patients with DM has grown. Proper management of these large volumes of data is expected to increase the quality of life of patients with DM [28]. Thus, AI may play a key role in the recognition of these systems as routine therapeutic aids for patients with DM.

Objective

Although several manuscripts have been published on the use of mobile and wearable technology for monitoring parameters that condition the development and evolution of DM, hereafter monitoring of DM-related parameters, this subject has not been systematically reviewed to the best of the authors' knowledge. Therefore, the goal of this study is to review the published literature on the use of mobile and wearable technology for the monitoring of DM-related parameters. Three specific research questions are defined to guide this study: (1) How are DM-related parameters studied using mobile and wearable technology? (2) How are the devices and sensors used to monitor DM-related parameters? and (3) What processing is given to the collected mobile and wearable data?

Methods

Overview

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [29] were followed to perform a systematic review of the literature on mobile and wearable sensing for the monitoring of DM-related parameters. Moreover, the Newcastle-Ottawa Scale (NOS) [30] was used to assess the quality of the studies. The specific methodology followed is described in the following sections.

Information Source

Studies were identified by searching electronic databases and scanning publications from a reference list of authors. The search was performed using 3 reference web-based citation databases: Web of Science (WoS), Scopus, and PubMed. The last search was performed on July 28, 2020. The queries used for the database search are listed in Textbox 1. The terms used in the queries and the combination thereof aim to match the title, abstract, or keywords of the manuscripts. In addition, we handsearched on ResearchGate for authors with skills and expertise related to the topics of interest by using the query "diabetes AND (wearable OR mobile)." On the basis of the list of retrieved authors, we identified their studies within the scope of wearable and mobile sensing in diabetes.

Textbox 1. Queries' results per database.

Scopus

- TITLE-ABS-KEY(diabetes AND ((sensor OR sensing OR accelerometer OR gyroscope OR "proximity sensor" OR "light sensor" OR pedometer OR barometer OR gps OR camera OR "humidity sensor" OR magnetometer OR compass OR microphone OR mic OR nfc OR Bluetooth OR Wi-Fi OR fingerprint OR sms OR "phone call" OR "call log") AND ((wearable OR "smart watch" OR smartwatch OR "fitness band" OR "flexible band" OR wristband OR "smart insole" OR bracelet) OR (mobile OR smartphone OR "smart phone" OR cellphone OR "cell phone" OR mobilephone OR "mobile phone"))))

Web of Science

- (ts = (diabetes AND ((sensor$ OR sensing OR accelerometer$ OR gyroscope$ OR "proximity sensor$" OR "light sensor$" OR pedometer$ OR barometer$ OR gps OR camera$ OR "humidity sensor$" OR magnetometer$ OR compass OR microphone$ OR mic OR nfc OR Bluetooth OR Wi-Fi OR fingerprint OR sms OR "phone call$" OR "phone$ call" OR "call log$") AND ((wearable$ OR "smart watch*" OR smartwatch* OR "fitness band$" OR "flexible band$" OR wristband$ OR "smart insole$" OR bracelet$) OR (mobile$ OR smartphone$ OR "smart phone$" OR cellphone$ OR "cell phone$" OR mobilephone$ OR "mobile phone$")))))

PubMed

- ((diabetes AND ((sensor OR sensing OR accelerometer OR gyroscope OR "proximity sensor" OR "light sensor" OR pedometer OR barometer OR gps OR camera OR "humidity sensor" OR magnetometer OR compass OR microphone OR mic OR nfc OR Bluetooth OR Wi-Fi OR fingerprint OR sms OR "phone call" OR "call log") AND ((wearable OR "smart watch" OR smartwatch OR "fitness band" OR "flexible band" OR wristband OR "smart insole" OR bracelet) OR (mobile OR smartphone OR "smart phone" OR cellphone OR "cell phone" OR mobilephone OR "mobile phone"))))[TitleAbstract])

Study Selection

Manuscripts resulting from the database search (WoS, Scopus, and PubMed) and the hand search (ResearchGate) were downloaded and merged, and duplicates were removed. A 2-stage process was applied for the analysis of the manuscripts. In the first stage, 2 of the authors (CR and OB) screened the manuscripts based on the eligibility criteria, using title and abstract. In the second stage, the same authors fully reviewed the manuscripts resulting from the first stage and selected those meeting the eligibility criteria. During both the initial screening and full-text screening for eligibility, the 2 authors processed all the papers independently and discussed their observations before making a definitive decision. In the event of disagreement, a third reviewer (CV) was assigned, and a final decision was made based on the majority vote.

Eligibility Criteria

Studies were included if related to DM and if data were collected using sensors from wearable devices or smartphones and transmitted wirelessly. Hence, studies that were not related to DM were directly excluded. Those related to DM but where data were not collected using wearables or smartphones or where data were not transmitted wirelessly were also excluded. The inclusion criteria for both disease and technology are explained below.

According to the considered disease, a manuscript was included if it focused exclusively on DM, meaning that the main clinical topic of the study was DM; it was related to DM complications, that is, the main clinical topic of the study was a complication (or several) resulting from DM; it studied DM in combination with another disease, in other words, the main clinical topic of the study was the relation between DM and another condition such as cardiovascular disease; and patients with DM were used as a case study, namely, a clinical solution for multiple conditions was proposed, but the evaluation was performed on patients with DM.

According to the technology, the definition of *wearable* used for the inclusion criteria was "electronic device with micro-controllers, that can be incorporated into clothing or worn on the body as implants or accessories" [31]. These devices can be either commercial, medical, or prototypes. The definition of *smartphone* for the inclusion criteria was the given by the Oxford dictionary: "a mobile phone that performs many of the functions of a computer, typically having a touchscreen interface, internet access, and an operating system capable of running downloaded apps." Moreover, both wearables and mobile devices must send the monitored data wirelessly to the storage endpoint to meet the inclusion criteria.

Studies meeting the disease and technology inclusion criteria were also excluded if they were oriented to the intervention without an actual monitoring of DM-related parameters; they were technology centered, namely, the solution was not applied to a clinical case study; the proposed solution was not tested; similar studies under a different title were already considered; and the manuscript was not available.

Only English manuscripts in engineering and computer science areas, of article or proceedings type, and published between January 2010 and July 2020 (both inclusive) were included.

Quality Assessment

The 9-point NOS was used to score the included manuscripts. Nonrandomized studies, including case-control and cohort studies, were independently scored by 2 authors (CR and JRR). Disagreements were discussed and resolved.

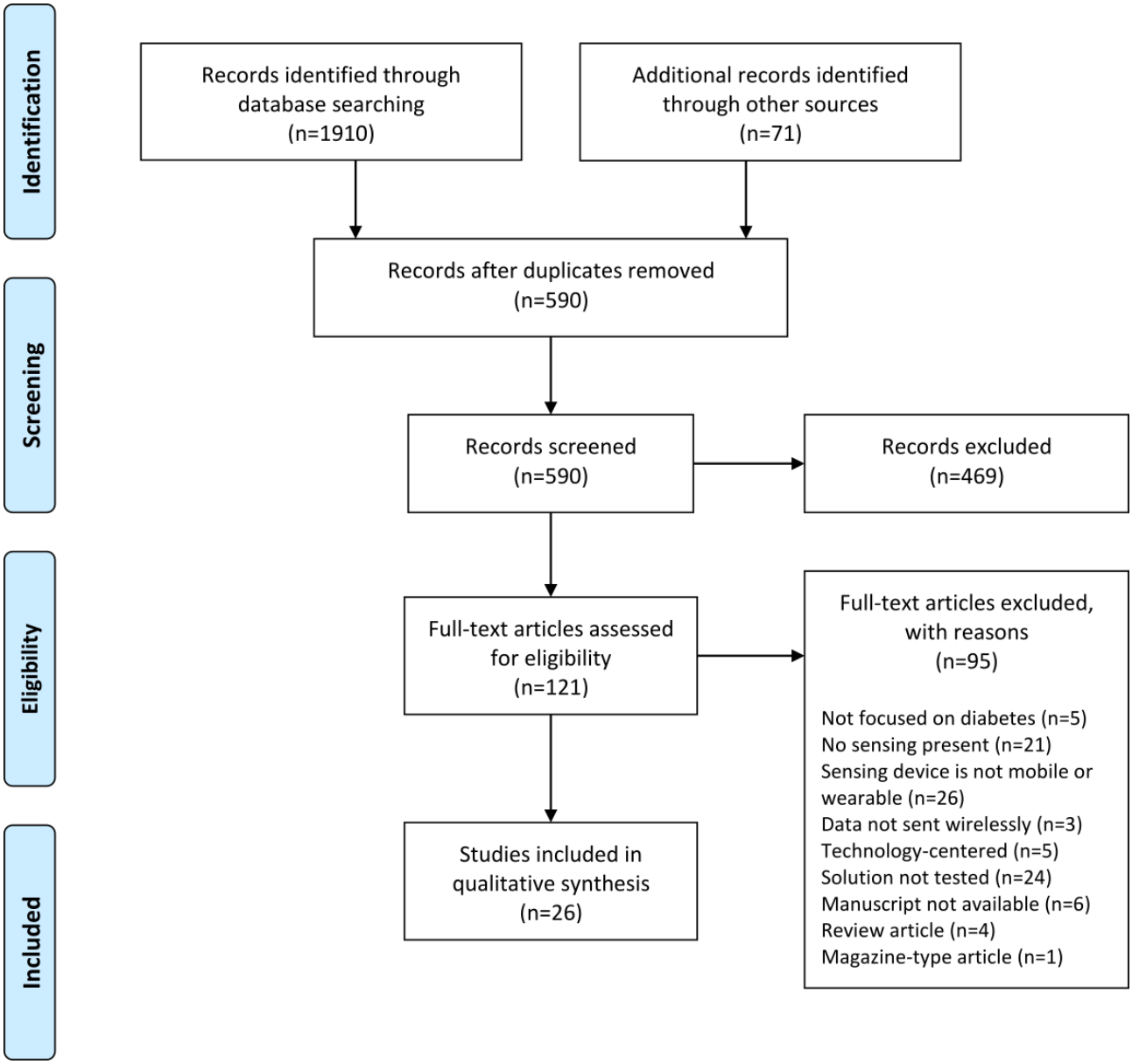
Results

Overview

The query used in the Scopus, WoS, and PubMed databases resulted in 960, 627, and 323 references, respectively. A total of 71 manuscripts were identified through other sources. After applying the PRISMA guidelines (Figure 1), 26 publications

were eventually included in the full review, 4 exclusive from Scopus [32-35]; 8 from Scopus and WoS [36-43]; 2 from Scopus and PubMed [44,45]; 1 from WoS and PubMed [46]; 9 from Scopus, WoS, and PubMed [47-55]; and 2 available in both Scopus and PubMed and other sources [56,57].

Figure 1. Search and selection of manuscripts using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.



The process used to achieve the 26 publications included in this review is as follows. A total of 590 original studies were obtained after merging the results from the databases and eliminating duplicates. Up to 79.5% (469/590) of the manuscripts were excluded after screening the title and abstract. Of the 469 manuscripts, 243 (51.8%) were not focused on DM, with a majority of papers where the term *diabetes* was part of the subject but most often just used as an exemplary use case; in 55 (11.7%) manuscripts, no sensing took place; in 4 (0.9%) of the manuscripts, the sensing device was not considered mobile or wearable; 97 (20.7%) manuscripts were mainly focused on interventions without an actual monitoring of DM-related parameters; 18 (3.9%) manuscripts were exclusively centered on the technology and did not have a clinical application; 1 (0.2%) study proposed a solution but simply at a conceptual level; 4 (0.9%) studies were found to be very similar to other considered studies despite being entitled differently; 11 (2.4%) were proceedings reviews; and finally 36 (7.7%) manuscripts were reviews or surveys.

The 121 manuscripts resulting from the previous screening were fully analyzed. A total of 78.5% (95/121) of the manuscripts were excluded following the same criteria mentioned above: 5% (5/95) studies were not focused on DM; in 22% (21/95) of the studies, no sensing was performed; in 27% (26/95) of the studies, the sensing device was not considered mobile or wearable; in 3% (3/95) of the studies, the data were not sent wirelessly; 5% (5/95) of the studies were centered on technology; 25% (24/95) of the studies were not properly tested, that is, some studies did not show enough scientific maturity in their tests and others did not involve patients with DM; 6% (6/95) of the manuscripts were not available; 4% (4/95) of the manuscripts were reviews or surveys; and finally, 1% (1/95) of

the manuscript was a magazine. As a result, 21.5% (26/121) of the manuscripts were selected to be reviewed in this study.

Of the 26 selected studies, 22 (85%) studies were assessed in terms of quality (Multimedia Appendix 1 [32-34,36-38,40-43,45-54,56,57]). A total of 19 (73%) studies were deemed as cohort studies [33,34,36-38,40-42,45-54,57]. Moreover, 12% (3/26) of studies were identified as case-control studies [32,43,56]. The remaining 15% (4/26) of studies were not identified as nonrandomized [35,39,44,55]; therefore, they were not included in the quality assessment. Of the assessed studies, 41% (9/22) were rated 1 star [36,37,40,43,46-49,51], 55% (12/22) were rated 2 stars [32-34,38,41,42,45,50,52-54,56], and only 5% (1/22) study was rated 4 stars [57]. None of the studies obtained any star for the comparability criteria. The ratings were generally low, which is nevertheless explained by the fact that the selected studies were not of an intervention type.

Some general statistics and quality indicators were obtained from the selected studies. Overall, 62% (16/26) manuscripts were published in journals, whereas (10/26) manuscripts were included in conference proceedings. Moreover, 23% (6/26) of the articles were published in journals ranked in quartile 1, 4% (1/26) in quartile 2, 8% (2/26) in quartile 3, and 4% (1/26) in quartile 4% (1/26) according to the Journal Citation Reports (WoS). In addition, 19% (5/26) of the articles were published in journals ranked in quartile 1, 23% (6/26) in quartile 2, 4% (1/26) in quartile 3, and 4% (1/26) in quartile 4 according to the SCImago Journal Rank (Scopus). It was observed that cross-national research teams were present on many occasions. These teams had a European member 14 times, had a member from the Unites States 11 times, and had an Asian member 8 times. The majority of the studies (21/26, 81%) were published between 2017 and 2019, as shown in Multimedia Appendix 2.

Close to half of the manuscripts (12/26, 46%) did not focus on a specific type of DM, almost a third of the studies (8/26, 31%) were related to T1D, 5 were associated with T2D (5/26, 19%), and 1 study dealt with GD (1/26, 4%). The majority of the studies (25/26, 96%) were some type of trial or longitudinal study, with sample sizes ranging from 1 to 100 participants with an approximate duration of up to 140 days. Patients with DM were involved in approximately two-third of these trials (17/26, 65%). More than half of the studies used only wearables to carry out their objectives (15/26, 58%), approximately one-fifth of the studies used only smartphones (5/26, 19%), and 6 studies used both wearables and smartphones in combination (6/26, 23%). The most common sensor used in the studies was the ACC (19/26, 73%).

In the following sections, we report the findings extracted from the analysis of all the reviewed studies. The dimensions for characterization are medical classification, DM type, research goals, devices, sensors, data processing, *privacy and security*, and *study's characteristics*. *Medical classification* refers to studies that focused exclusively on DM, related to DM complications, studied DM in combination with another disease, or used patients with DM as a case study. *DM type* can be T1D, T2D, GD, or not specified. *Research goals* refer to activity recognition, diagnosis or prediction of the onset or evolution of DM, finding associations among DM-related variables, or simply measuring DM-related parameters. *Devices* can be *wearable*, *smartphone*, *wearable and smartphone*. *Sensors* refer to ACC, glucose monitor (GM), HR monitor, etc. *Data processing* is categorized into statistical analysis, machine learning (ML), ontologies, or none. *Privacy and security* refer to mechanisms used to preserve users' anonymity, ethical aspects, and data protection. *Study's characteristics* refer to size, length, and subjects' characteristics.

Medical Classification, DM Type, and Research Goals

More than half of the selected studies, namely, 62% (16/26) of manuscripts were exclusively focused on DM [32,34-40,45-47,50,51,53-55]. Other studies addressed some complications that DM can lead to, specifically, diabetic foot ulcer (DFU) in 15% (4/26) of articles [33,48,56,57] and diabetic peripheral neuropathy (DPN) in 4% (1/26) of studies [56]. Three additional studies (12%) investigated DM in conjunction with other diseases that were not a complication of DM itself but were closely related to it. Nguyen et al [41] considered cardiovascular diseases besides DM, whereas in the study by Sarda et al [52], the association between DM and depression was explored. Sevil et al [42] found an association between acute psychological stress and glucose dynamics in patients with T1D. Some of the screened manuscripts were not focused on a specific disease, but their results can be applied to different health domains, such as obesity, nutrition, hypertension, and DM. Furthermore, 12% (3/26) of these studies were included in the final set of selected articles because they used patients with DM as a case study [43,44,49].

According to the DM type, 4% (1/26) of studies were related to GD [49], 31% (8/26) of studies were related to T1D [32,38,42,45,50,51,53,54], 19% (5/26) of studies were related to T2D [34,39,43,46,55], and 46% (12/26) of studies did not mention any specific type of DM [33,35-37,40,41,44,47,48,52,56,57].

As for the research goals of the selected studies, 15% (4/26) of manuscripts aimed to detect physical activity patterns relevant to patients with DM, such as walking, running, or sleeping [36,37,41,47]. Three studies [36,37,47] focused on recognizing eating, exercising, and sleeping activities by using body ACC and HR collected via a smartphone and a chest strap. In addition, 2 studies [37,47] collected sound, location, velocity, and respiration rate (RR) for a similar purpose. Nguyen et al [41] also used a chest band and ambient sensors to measure ACC, continuous GL, body temperature, room temperature, or humidity data, which are used for fall detection and remote monitoring of DM-related parameters.

Overall, 23% (6/26) of manuscripts [34,35,38,39,48,53] focused on diagnosing or predicting affections related to DM from sensor data. Calbimonte et al [38] predicted glycemic events by collecting ECG, RR, and ACC from a chest strap. Fraiwan et al [48] created a mobile thermal imaging system to indicate the potential development of a DFU. Reddy et al [39] aimed to create a classification system to identify an individual's diabetic status (healthy or diabetic) using photoplethysmogram data collected from both a smartphone and pulse oximeter. Ramazi et al [34] predicted the progression of T2D by collecting the

physical activity and GL from a wristband and a CGM. Garcia et al [35] diagnosed DM from facial images captured using a smartphone. Finally, Rodriguez-Rodriguez et al [53] predicted GL using the CGM data alone.

A total of 38% (10/26) of studies were aimed at finding an association between variables or diseases related to DM [32,42,44,45,50-52,54,56,57]. Najafi et al [44] found some links among biometric variables, postural and balance control, from the quaternion data collected via a hip-worn strap prototype. Turksoy et al [50] determined which variable was the most useful for inclusion in a future artificial pancreas using data such as ACC, GL, and HR from various wearable devices such as an armband, a CGM, and a smartwatch. Faccioli et al [45] studied the relation between physiological information on physical exercise and glucose models, mainly from ACC and GL data sensed via an activity tracker and CGM. Similarly, Merickel et al [32] found a relation between the driver state and the health and physiological state of patients with T1D using a wristwatch and a CGM to obtain primarily ACC, GL, and HR. Likewise, Sarda et al [52] analyzed the relation between some variables sensed by a smartphone, such as mobility, sleep patterns, and location, and the symptoms of depression with DM. Sevil et al [42] found an association between acute psychological stress and glucose dynamics. Similarly, Sanz et al [54] sought to find associations between different signals provided by 3 different wearable devices and the accuracy of a CGM device during aerobic exercise. Two studies found associations between diseases. On the one hand, Grewal et al [56] found a relation between DPN, with no other complications, and DPN combined with DFU using spatiotemporal parameters from a lower limb band. On the other hand, Razjouyan et al [57] found a relation between physiological stress and DFU healing speed via the monitoring of ACC, HR, and RR from a chest strap. Finally, Groat et al [51] attempted to correlate biometric and physiological variables sensed by a CGM and a wristband through manual entries on a mobile app.

Finally, 23% (6/26) of manuscripts had the principal goal of simply measuring data relevant to DM [33,40,43,46,49,55]. McLean et al [49] sensed physical proximity, physical activity, sedentary behavior, location, and orientation from a smartphone. Bartolic et al [40] primarily used a wristband and a smartphone to monitor daily activity, sleep duration, and calorie consumption, among others. McMillan et al [46] collected physical activity data, sedentary behavior, and continuous GL using a thigh-worn wearable and a flash glucose monitor (FGM). Rescio et al [33] sensed the temperature and pressure of the plantar foot with a smart insole. Zherebtsov et al [43] quantified changes in the microcirculatory blood flow in tissues using a wristband. Whelan et al [55] measured the usage, feasibility, and acceptability of behavioral and physiological self-monitoring technologies with 2 different wristbands and an FGM.

All details about medical classification, DM type, and research goals of the reviewed manuscripts are provided in Table 1.

Table 1. Summary of medical topics, diabetes types, and research goals, ordered by the year of publication.

Manuscript	Medical classification	Diabetes type	Research goals
Najafi et al (2010) [44]	Patients with DM[a] used as case study	Not specified	Find association between posture and balance control among patients with different DM complications
Grewal et al (2013) [56]	DM complication: DPN[b] and DPN with diabetic foot	Not specified	Find association between DPN and DFU[c] for gait
Luštrek et al (2014) [36]	Focused on DM	Not specified	Activity recognition of walking, running, cycling, lying, sitting, and standing
Luštrek et al (2015) [37]	Focused on DM	Not specified	Activity recognition of sleeping, home chores, home leisure, eating, and exercising
Cvetković et al (2016) [47]	Focused on DM	Not specified	Activity recognition of working, eating, exercising, and home activities
Calbimonte et al (2017) [38]	Focused on DM	T1D[d]	Predict glycemic events
Fraiwan et al (2017) [48]	DM complication: diabetic foot	Not specified	Diagnose development of DFU
McLean et al (2017) [49]	Patients with DM used as case study	GD[e]	Measure physical proximity, physical activity, and magnetic field strength
Razjouyan et al (2017) [57]	DM complication: diabetic foot	Not specified	Find association between physiological stress response and healing speed among outpatients with active DFU
Reddy et al (2017) [39]	Focused on DM	T2D[f]	Diagnose individual's diabetic status
Turksoy et al (2017) [50]	Focused on DM	T1D	Find association between biometric variables and changes in glucose concentration
Bartolic et al (2018) [40]	Focused on DM	Not specified	Measure GL[g], insulin dosage, physical activity, daily movement, and sleep duration and quality
Faccioli et al (2018) [45]	Focused on DM	T1D	Find association between glucose prediction models' performance
Groat et al (2018) [51]	Focused on DM	T1D	Find association between exercise behavior data with the rate of change in GL
McMillan et al (2018) [46]	Focused on DM	T2D	Measure combined GL data, physical activity, and sedentary behavior
Merickel et al (2018) [32]	Focused on DM	T1D	Find association between pattern of glucose and at-risk pattern of vehicle acceleration behavior
Nguyen Gia et al (2019) [41]	DM in conjunction with other diseases: DM+cardiovascular disease	Not specified	Activity recognition of fall detection and remote health monitoring
Rescio et al (2019) [33]	DM complication: diabetic foot	Not specified	Measure temperature and pressure of the plantar foot
Sarda et al (2019) [52]	DM in conjunction with other diseases: DM+depression	Not specified	Find association between smartphone-sensing parameters and symptoms of depression
Ramazi et al (2019) [34]	Focused on DM	T2D	Predict the progression of T2D
Garcia et al (2019) [35]	Focused on DM	Not specified	Diagnose DM from facial images
Sevil et al (2019) [42]	DM in conjunction with other diseases: DM+acute psychological stress	T1D	Find the association between acute psychological stress and the glucose dynamics
Zherebtsov et al (2019) [43]	Patients with DM used as case study	T2D	Measure the changes in the microcirculatory blood flow of healthy patients and patients with T2D
Rodriguez-Rodriguez et al (2019) [53]	Focused on DM	T1D	Predict blood GL for T1D with limited computational and storage capabilities using only CGM[h] data
Sanz et al (2019) [54]	Focused on DM	T1D	Find the association between different signals provided by 3 different wearables devices and the accuracy of a CGM device during aerobic exercises

Manuscript	Medical classification	Diabetes type	Research goals
Whelan et al (2019) [55]	Focused on DM	T2D	Measure the use, feasibility, and acceptability of behavioral and physiological self-monitoring technologies in individuals at risk of developing T2D

[a]DM: diabetes mellitus.

[b]DPN: diabetic peripheral neuropathy.

[c]DFU: diabetic foot ulcer.

[d]T1D: type 1 diabetes.

[e]GD: gestational diabetes.

[f]T2D: type 2 diabetes.

[g]GL: glucose level.

[h]CGM: continuous glucose monitor.

Devices and Sensors

A total of 58% (15/26) of studies used wearable devices for monitoring tasks [32-34,38,41-44,46,50,53-57], 3 of which were research prototypes, namely, a hip-worn strap [44], a chest strap [41], and a smart insole [33]. Of the remaining 12 studies, 5 (42%) used wearable of a commercial type; those were chest straps (Bioharness 3, Zephyr) [38,57], a lower limb band (LEGSys, BioSensics LLC) [56], a wristband (AMT-LAZMA 1, Aston Medical Technology Ltd) [43], and an FGM (FreeStyle Libre, Abbott Diabetes Care) [53]. In 23% (6/26) of studies, the authors used various commercial wearable devices in combination: an armband (SenseWear, BodyMedia), 2 CGMs (Guardian Real-Time, Medtronic, and Dexcom G4 Platinum, Dexcom, Inc), a chest strap (Bioharness 3, Zephyr), and a smartwatch (Mio Alpha, MIO Global) [50]; a thigh-worn (activPAL, PAL Technologies Ltd) and a FGM (FreeStyle Libre, Abbott Diabetes Care) [46]; a wristwatch (model not specified) and a CGM (Dexcom G4 Platinum, Dexcom, Inc) [32]; a wristband (ActiGraph, model not specified, ActiGraph LLC) and a CGM (Dexcom G4 Platinum, Dexcom, Inc) [34]; a wristband (Empatica E4, Empatica Inc) and a CGM (Dexcom G5, Dexcom, Inc) [42]; 3 wristbands (Fitbit Charge HR, Fitbit, Inc; Microsoft Band 2, Microsoft Corporation; and Polar RCX3, Polar Electro) and a CGM (Enlite-2, Medtronic Minimed) [54]; and 2 wristbands (Fitbit Charge 2, Fitbit, Inc, and ActiGraph wGT3x-BT, ActiGraph LLC) and a FGM (FreeStyle Libre, Abbott Diabetes Care) [55].

Smartphones were used in 19% (5/26) of studies for sensing purposes: a Samsung Galaxy S6 Edge Plus (Samsung) [48], a Nexus 5 (LG Corporation & Google, LLC) [39], an iPhone 7 (Apple Inc) [35], and no specific details were found for the other 2 studies [49,52]. A total of 6 studies used both wearables and smartphones to acquire data [36,37,40,45,47,51]. In 3 studies [36,37,47], a smartphone (model not specified) and a chest strap (model not specified) were used. Bartolic et al [40] used a smartphone (model not specified) with a wristband (Fitbit fitness bracelet, Fitbit, Inc). Faccioli et al [45] used a smartphone (model not specified), an activity tracker (model not specified), and a CGM (model not specified). Groat et al [51] used a smartphone (model not specified), a wristband (Fitbit Charge HR, Fitbit, Inc), and a CGM (Enlite, Medtronic) to perform the monitoring task.

A total of 30 different types of sensors were used to acquire data from the reviewed studies. The ACC stands out as the sensor most widely used across studies, namely, in 73% (19/26) of studies [32,34,36-38,40-42,44-47,49,50,52,54-57]. This sensor has been primarily used for the automatic recognition of activities such as motion, walking, running, exercise, cycling, standing, sitting, sleep, step count, and fall detection. The second most used sensor is the GM, which was used to monitor glucose in 46% (12/26) of the most recently published studies [32,34,40-42,45,46,50,51,53-55]. The subject's glucose was measured using different types of GM sensors: continuously with a CGM [32,34,41,42,45,50,51,54], with an FGM [46,53,55], and manually with a Bluetooth glucometer [40] and with a glucose and L-lactate analyzer [54]. Moreover, HR monitor was included in 27% (7/26) studies [32,36,40,50,51,54,55] to measure the HR of the subjects and infer mainly exercise intensity, calorie consumption, and activity recognition. Some apps can also be considered as sensors when used to encode data manually. For example, Groat et al [51] used an app to be operated by the subjects involved in the trials to self-track exercise behavior and rate of change in GLs.

All details about the devices and sensors used in the reviewed manuscripts are provided in Table 2. A summary of the sensor use for each study is available in Multimedia Appendix 3 [32-57].

Table 2. Summary of devices used, ordered by the year of publication.

Studies and devices	Devices, sensors, measurement	Purpose
Najafi et al (2010) [44]	Wearable: hip-worn strap (*prototype*[a]) • Triaxial ACC[b] (quaternions) • Triaxial gyroscope (quaternions) • Triaxial magnetometer (quaternions) Complementary: pressure platform (*Emed-x system, Novel Inc*) • Pressure sensor (area of sway)	Recognize the motion of ankle and hip joints in 3 dimensions
Grewal et al (2013) [56]	Wearable: lower limb band (*LEGSys, BioSensics LLC*) • ACC (acceleration) • Gyroscope (angular velocity)	Gait detection
Luštrek et al (2014) [36]	Smartphone (*model not specified*) • ACC (acceleration) Wearable: chest strap (*model not specified*) • ACC (acceleration) • HRM[c] (HR[d])	Smartphone location detection; activity recognition: walking, running, cycling, lying, sitting, and standing; and energy expenditure estimation
Luštrek et al (2015) [37]	Smartphone (*model not specified*) • ACC (acceleration, location, HR, and RR[e]) • Microphone (sound) • GPS (location and velocity) • Wi-Fi (location) Wearable: chest strap (*model not specified*) • ACC (acceleration) • ECG[f] (HR and RR)	Activity recognition: sleep, exercise, work, transport, eating, home, and outdoor
Cvetković et al (2016) [47]	Smartphone (*model not specified*) • ACC (acceleration) • Microphone (sound) • GPS (location and velocity) • Wi-Fi (location) Wearable: chest strap (*model not specified*) • ACC (acceleration) • ECG (HR and RR)	Activity recognition: sleep, exercise, work, transport, eating, home, and out
Calbimonte et al (2017) [38]	Wearable: chest strap (*Bioharness 3, Zephyr*) • ACC (acceleration) • ECG (HB[g] fiducial points location, ST[h] segment shape, QTc[i] interval, HR, and RR)	Generate 2 semantic models: physiological and energy expenditure for classifying hypoglycemic events
Fraiwan et al (2017) [48]	Smartphone (*Samsung Galaxy S6 Edge Plus, Samsung*) Complementary: infrared thermal camera (*FLIR ONE, FLIR Systems, Inc*) • Infrared sensor (thermal images) • Camera (standard image)	Recognize change of temperature on the feet
McLean et al (2017) [49]	Smartphone (*model not specified*) • ACCs (acceleration) • GPS (location) • Wi-Fi (location) • Camera (photo) • Magnetometer (magnetic field strength) • Bluetooth (physical proximity)	Quantify physical proximity, sedentary behavior, vehicle use, and location
Razjouyan et al (2017) [57]	Wearable: chest strap (*Bioharness 3, Zephyr*) • ACC (acceleration) • ECG (HR, RR, and core body temperature)	Detection of physiological stress of the patient

Studies and devices	Devices, sensors, measurement	Purpose
Reddy et al (2017) [39]	Smartphone (*Nexus 5, LG Corporation and Google, LLC*) • Camera and flash (PPG[j]) Complementary: peripheral pulse oximeter (*model not specified*) • Pulse oximeter (PPG)	Discriminate between diabetic and healthy individuals
Turksoy et al (2017) [50]	Wearable: armband (*SenseWear, BodyMedia*) • ACC (acceleration) • Thermometer (skin temperature and near-body temperature) • Galvanometer (galvanic skin response) • Heat flux (rate of heat dissipating from the body) Wearable: CGM (*Guardian Real-Time, Medtronic*) • GM[k] (GL[l]) Wearable: CGM (Dexcom *G4 Platinum, Dexcom, Inc*) • GM (GL) Wearable: chest strap (*Bioharness 3, Zephyr*) • ACC (acceleration) • ECG (HR) Wearable: smartwatch (*Mio Alpha, MIO Global*) • HRM (HR) Complementary: open-circuit spirometry metabolic cart system (*True One, Parvo Medics*) • Expired gases (O2 and CO2)	Find a correlation between biometric changes and glucose concentrations during exercise
Bartolic et al (2018) [40]	Smartphone (*model not specified*) • App (insulin doses) Wearable: wristband (*Fitbit fitness bracelet, Fitbit, Inc*) • ACC (acceleration) • HRM (HR) Complementary: glucometer (*Contour Next One, Ascensia Diabetes Care*) • GM (GL)	Quantify physical activity, daily movement, sleep duration and quality, calorie consumption, insulin dosages, and continuous GL
Faccioli et al (2018) [45]	Smartphone (*model not specified*) • App (carbohydrates count) Wearable: activity tracker (*model not specified*) • ACC (acceleration) Wearable: CGM (*model not specified*) • GM (continuous GL)	Quantify step count, continuous GL, and carbohydrate intake
Groat et al (2018) [51]	Smartphone (*model not specified*) • App (exercise behavior) Wearable: wristband (*Fitbit Charge HR, Fitbit, Inc*) • HRM (HR) Wearable: CGM (*Enlite, Medtronic*) • GM (continuous GL)	Quantify exercise behavior measured via a wristband and an app to compare with the rate of change in GL recorded by a CGM
McMillan et al (2018) [46]	Wearable: thigh-worn (*activPAL, PAL Technologies Ltd)* • ACC (acceleration) • Inclinometer (acceleration) Wearable: FGM (*FreeStyle Libre, Abbott Diabetes Care*) • GM (continuous GL)	Quantify step count, cadence and postural transitions and energy expenditure estimates, and continuous GL
Merickel et al (2018) [32]		

Studies and devices	Devices, sensors, measurement	Purpose
	Wearable: wristwatch (*model not specified*) • ACC (acceleration) • HRM (HR) Wearable: CGM (Dexcom *G4 Platinum, Dexcom, Inc*) • GM (continuous GL) Complementary: vehicle sensor instrumentation packages (*model not specified*) • Camera (video) • GPS (vehicle acceleration and speed) • OBD[m] sensor (vehicle acceleration and speed)	Compare the driving behavior from drivers with and without T1D[n]
Nguyen Gia et al (2019) [41]	Wearable: chest strap (*prototype*) • ACC (acceleration) • Gyroscope (angular velocity) • Magnetometer (magnetic field) • ECG (QT[o] intervals and HR) • GM (continuous GL) • Thermometer (body temperature) Complementary: ambient sensors (*prototype*) • Ambient sensor (room temperature, humidity, and air quality)	Monitor DM[p] and ECG, and report abnormalities: fall, very low or high GL, and abnormal HR in real time without interfering with the patient's daily activities
Rescio et al (2019) [33]	Wearable: smart insole (*prototype*) • Infrared thermometer (plantar temperature) • Pressure sensor (pressure)	Monitor temperature and pressure of the plantar foot
Sarda et al (2019) [52]	Smartphone (*model not specified*) • ACC (acceleration) • Call logs (communication) • GPS (location) • Ambient light sensor (ambient light)	Activity recognition: mobility, sleep, and social interaction
Ramazi et al (2019) [34]	Wearable: CGM (*DexcomG4 Platinum, Dexcom, Inc*) • GM (continuous GL) Wearable: wristband (*model not specified; ActiGraph, ActiGraph LLC*) • ACC (acceleration)	Quantify GL, traveled steps, and physical activity: sitting, standing, and lying
Garcia et al (2019) [35]	Smartphone (*iPhone 7, Apple Inc*) • Camera (standard image)	Capture facial images
Sevil et al (2019) [42]	Wearable: wristband (*Empatica E4, Empatica Inc*) • ACC (acceleration) • PPG (blood volume pulse) • Galvanometer (galvanic skin response) • Infrared thermopile (skin temperature) Wearable: CGM (*Dexcom G5, Dexcom, Inc*) • GM (continuous GL)	Estimate acute psychological stress effect index and GL
Zherebtsov et al (2019) [43]	Wearable: wristband (*AMT-LAZMA 1, Aston Medical Technology Ltd*) • Laser Doppler flowmetry (Doppler shift)	Quantify changes in the microcirculatory blood flow in tissues
Rodriguez-Rodriguez et al (2019) [53]	Wearable: FGM (*FreeStyle Libre, Abbott Diabetes Care*) • GM (continuous GL)	Quantify GL for the creation of a database for further processing by the prediction models
Sanz et al (2019) [54]		Quantify number of steps walked, number of floors of stairs climbed, exercise intensity, calories burned, and skin electrodermal activity

Studies and devices	Devices, sensors, measurement	Purpose
	Wearable: wristband (*Fitbit Charge HR, Fitbit, Inc*) • ACC (acceleration) • HRM (HR) • Altimeter (altitude) Wearable: wristband (*Microsoft Band 2, Microsoft Corporation*) • HRM (HR) • Skin temperature (skin temperature) • Galvanometer (galvanic skin response) • ACC (acceleration) Wearable: wristband (*Polar RCX3, Polar Electro*) • HRM (HR) Wearable: CGM (*Enlite-2, Medtronic Minimed*) • GM (continuous GL) Complementary: glucose and L-lactate analyzer (*YSI 2300 Stat Plus Glucose Analyzer, YSI Incorporated Life Sciences*) • GM (GL)	
Whelan et al (2019) [55]	Wearable: wristband (*Fitbit Charge 2, Fitbit, Inc*) • ACC (acceleration) • Altimeter (altitude) • HRM (HR) Wearable: wristband (*ActiGraph wGT3x-BT, ActiGraph LLC*) • ACC (acceleration) Wearable: FGM (*FreeStyle Libre, Abbott Diabetes Care*) • GM (continuous GL)	Quantify number of steps walked, distance traveled, HR, calories expended, flights of stairs climbed, and GL

[a]Text in italic represents model and company of each devices in that order. In the cases of no specification on the correspondent manuscript "Model not specified" it is stated.

[b]ACC: accelerometer.

[c]HRM: heart rate monitor.

[d]HR: heart rate.

[e]RR: respiration rate.

[f]ECG: electrocardiogram.

[g]HB: heartbeat.

[h]ST: electrocardiogram measurement ST interval.

[i]QTc: corrected electrocardiogram measurement QT interval.

[j]PPG: photoplethysmogram.

[k]GM: glucose monitor.

[l]GL: glucose level.

[m]OBD: on-board diagnostics device.

[n]T1D: type 1 diabetes.

[o]QT: electrocardiogram measurement QT interval.

[p]DM: diabetes mellitus.

Data Processing

Except for 12% (3/26) of studies [33,46,49], most manuscripts included some type of processing to the collected data. In 42% (11/26) of studies, the processing mainly consisted of statistical analysis [32,40,43-45,50,51,54-57]; in 23% (6/26) of studies, the authors used ML techniques [35-37,39,41,47], and in 15% (4/26) of studies [34,42,52,53], both statistical analysis and ML were used. Fraiwan et al [48] used image processing to detect areas with varying body temperature, and Calbimonte et al [38] used ontologies for diagnosing glycemic events.

The studies without specific data processing were rather oriented to simply collect data. ML was primarily used in the reviewed investigations for activity recognition and diagnosis tasks. Statistical analyses were predominantly aimed at finding associations between behavioral and physiological variables with DM conditions.

Some authors created their own algorithm to attain the study objective. Cvetković et al [47] developed 4 new methods to recognize human activity: person-dependent, person-independent, person-independent with person-specific data, and person-independent and person-specific models

combined with heuristics and multiclassifier adaptive training. Merickel et al [32] developed their own procedure to eliminate the spurious values of GLs and to discretize them. Nguyen et al [41] created new algorithms for HR and the QT interval extraction from the ECG for activity status categorization and for fall detection. Finally, Ramazi et al [34] developed a new algorithm for different sensor signal synchronization.

Some studies used additional data, not collected passively from mobile or wearable devices, to achieve their goal. One example is the study by Razjouyan et al [57], where the authors collected clinical information such as depression scale, numeric pain scale, and glycated hemoglobin level by using questionnaires, in addition to demographic information. Similarly, Ramazi et al [34] used clinical information, such as triglycerides, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and very low-density lipoprotein cholesterol, and demographic information to improve the prediction of T2D progression in patients. Furthermore, in the study by Turksoy et al [50], the diet and physical activity of the subjects were documented manually throughout the remainder of the study period. Sarda et al [52] collected sociodemographic information, such as gender, marital status, occupation, or education, to perform descriptive analysis to understand the societal representation of the participants. Finally, Sevil et al [42] recorded, besides sensor data, nutrition data (carbohydrate ingestion times, content, and amount); demographic information; and other information, such as mood, sleep information, and feelings of anxiety or depression, using questionnaires.

All details about data processing for the reviewed manuscripts are provided in Table 3.

Table 3. Summary of processing techniques, privacy, and security ordered by the year of publication.

Manuscript	Statistical methods	Machine learning methods	Privacy	Security
Najafi et al (2010) [44]	Pearson correlation coefficient, paired *t* test, and intraclass correlation coefficient	None	Study approved by the local ethics committee	Not described
Grewal et al (2013) [56]	Statistical fluctuation, SD, and coefficient of variation	None	Study approved by the local ethics committee. All participants signed an informed consent form before participating in the study	Not described
Luštrek et al (2014) [36]	None	RF[a] and support vector regression algorithm	Not described	Not described
Luštrek et al (2015) [37]	None	Spectral centroid, zerocrossing, mel frequency cepstral coefficient, linear predictive coding, and method of moments values of the sound signals; clustering of Wi-Fi and GPS data; new algorithm of acceleration data; naive bayes, logistic regression, SVM[b], RF, RIPPER[c], adaboost, and bagging for activity recognition tasks; and event calculus for interpreting recognized activities	Sound from the smartphone microphone is recorded in fractions of 100 ms per second	Not described
Cvetković et al (2016) [47]	None	Spectral centroid, zerocrossing, mel frequency cepstral coefficient, linear predictive coding, and method of moments values of the sound signals; clustering of Wi-Fi and GPS data; new algorithm of acceleration data; 5 new algorithms for activity recognition task; and symbolic rules to refine confused predictions	Sound from the smartphone microphone is recorded in fractions of 100 ms per second	Not described
Calbimonte et al (2017) [38]	None	Normalized least mean squares, ontology, and RDF[d] stream processing engine (CQELS[e] continuous evaluation)	Not described	Not described
Fraiwan et al (2017) [48]	None	Otsu thresholding technique and point-to-point mean difference technique	Authors referred to "Ethics approval and consent to participate" as "Not applicable"	Not described
McLean et al (2017) [49]	None	None	Not described	Data are stored on the phone and uploaded in an encrypted form
Razjouyan et al (2017) [57]	Analysis of variance, root mean square of successive R-wave to R-wave intervals, power spectrum density of time series representing R-wave to-R-wave intervals, receiver operating characteristic, and area under the curve	None	Not described	Not described
Reddy et al (2017) [39]	None	SVM, artificial neural network, and classification and regression trees	Not described	Not described
Turksoy et al (2017) [50]	Partial least squares, regression, and variable importance in projection	None	Not described	Not described
Bartolic et al (2018) [40]	Trading view and minimum and maximum values	None	Not described	Not described

Manuscript	Statistical methods	Machine learning methods	Privacy	Security
Faccioli et al (2018) [45]	Black-box linear model, prediction error method, coefficient of determination, and RMSE[f]	None	The trial study and all experimental procedures were approved by the institution's ethical review board	Not described
Groat et al (2018) [51]	Cohen κ	None	Study approved by the local institutional review board	Not described
McMillan et al (2018) [46]	None	None	Not described	Not described
Merickel et al (2018) [32]	Their own procedures and β regression model	None	All subjects gave informed consent to study participation according to the University of Nebraska Medical Center's institutional review board's protocols	Not described
Nguyen Gia et al (2019) [41]	None	Heart rate and the QT[g] interval extraction, activity status categorization, and fall detection	Not described	Lightweight cryptography
Rescio et al (2019) [33]	None	None	Not described	Not described
Sarda et al (2019) [52]	Descriptive analysis and univariate analysis	SVM, RF, adaboost, extreme gradient boosting, and cross-validation	Not described	All transmissions were in an encrypted form using the HTTPS[h] secure sockets layer protocol
Ramazi et al (2019) [34]	RMSE	New algorithm for different sensor signal synchronization and long short-term memory deep neural network	Study approved by the local institutional review board	Not described
Garcia et al (2019) [35]	None	KNN[i] and SVM	Not described	Not described
Sevil et al (2019) [42]	Mean, SD, kurtosis, and mean absolute error	SVM, KNN, linear discriminant, decision tree, and logistic regression	Study approved by the local institutional review board	Not described
Zherebtsov et al (2019) [43]	Statistical significance	None	Study approved by the local institutional review board. Each volunteer gave a voluntary informed written consent to participate in the experiment	Not described
Rodriguez-Rodriguez et al (2019) [53]	RMSE	Autoregressive integrated moving average, RF, and SVM	Study conducted in accordance with the Helsinki Declaration. Study approved by the local ethics committee. Data storage complied with the stricter data protection rules for protecting personal information. All participants were fully informed about the purpose of the experiment and provided written informed consent and assent according to the national regulations	Not described

Manuscript	Statistical methods	Machine learning methods	Privacy	Security
Sanz et al (2019) [54]	Median, linear regression, and cross-validation	None	Study approved by the local ethics committee	Not described
Whelan et al (2019) [55]	Mean, SD, and frequency	None	All participants provided written informed consent. Study approved by the local ethics advisory committee	Not described

[a]RF: random forest.

[b]SVM: support vector machine.

[c]RIPPER: repeated incremental pruning to produce error reduction.

[d]RDF: resource description framework.

[e]CQELS: continuous query evaluation over linked stream.

[f]RMSE: root mean square error.

[g]QT: electrocardiogram measurement QT interval.

[h]HTTPS: Hypertext Transfer Protocol Secure.

[i]KNN: k-nearest neighbors.

Privacy and Security

A total of 62% (16/26) of studies addressed privacy or security issues [32,34,37,41-45,47,49,51-56]. Of the 13 studies dealt with privacy aspects to some extent: 10 (77%) studies were approved by a local ethics committee or review boards [34,42-45,51,53-56]; in the study by Merickel et al [32], the authors included informed consent for participants; and in 2 (15%) studies [37,47], the sound from the smartphone microphone was downsampled to 100 ms out of every second to preserve the user's privacy. Finally, the study by Fraiwan et al [48] was not included in the above 16 studies because it was mentioned that the intervention of an ethics committee was not applicable.

Overall, 19% (3/16) of studies considered security aspects. McLean et al [49] stored the data on the phone and then uploaded the data to a server in an encrypted manner. Nguyen et al [41] used lightweight cryptography to deal with security in the mobile monitoring system, and Sarda et al [52] stated that all transmissions were in an encrypted form using the Hypertext Transfer Protocol Secure (HTTPS) secure sockets layer protocol.

The remaining 35% (19/26) of studies did not mention anything about privacy or security [33,35,36,38-40,46,50,57].

All details about *privacy and security* of the reviewed manuscripts are provided in Table 3.

Study's Characteristics

The number of participants involved differed significantly among the studies. The average number of participants was 29 (SD 28.2), calculated from 88% (23/26) of studies that indicated the number of participants [32-37,39,41-47,49-57]. The minimum sample size was 1 subject [46], and the maximum sample size was 100 subjects [35,39]. Furthermore, 12% (3/26) of papers did not specify this number [32,38,48]. The duration of the test phase was standardized to days, namely, an average of 21.5 (SD 35.1) days of duration as computed from the 58% (15/26) of studies that specified this value [32,34,37,39,42,43,45-47,50-53,55,57]. One day was the minimum duration of the study [42,46], and 140 days was the maximum duration of the study [52]. In 13% (2/15) studies [39,43], the duration remained in the order of minutes for a better understanding, and 42% (11/26) of studies did not specify this value at all [33,35,36,38,40,41,44,48,49,54,56].

As for the health status distribution for the 26 studies, 11 (42%) studies involved patients with DM [34,42,45,46,49-54,57], 5 (19%) included healthy subjects [36,41,47,48,55], and 6 (23%) involved both healthy subjects and patients with DM [32,35,39,43,44,56]. Moreover, 12% (3/26) of studies did not precisely describe the health status of the subjects [33,37,40], and 4% (1/26) of studies used an existing data set to test their solution [38].

A total of 31% (8/26) of studies involved subjects of both genders [35,37,47,50-53,55], and 4% (1/26) of studies only involved a male [46]. The remaining 65% (17/26) of studies did not specify the gender of the participants [32-34,36,38-45,48,49,54,56,57]. With respect to the age of the involved subjects, 65% (17/26) of studies provided this value [32-35,39,41,43,46,47,50-57], resulting in an average age of 44.5 (SD 12.2) years. The other 35% (9/26) of studies did not mention any age distribution [36-38,40,42,44,45,48,49].

All details about the *study's characteristics* for the reviewed manuscripts are provided in Table 4.

Table 4. Summary of study topic ordered by the year of publication.

Manuscript	Sample size	Sample type	Duration
Najafi et al (2010) [44]	38	17 diabetic and 21 healthy; gender undefined; and age undefined	Not described
Grewal et al (2013) [56]	39	31 diabetic and 8 healthy; gender undefined; and aged 56.9 (SD 8.2) years	Not described
Luštrek et al (2014) [36]	10	0 diabetic and 10 healthy; gender undefined; and age undefined	Not described
Luštrek et al (2015) [37]	5	Health status undefined; 1 female and 4 males; and age undefined	14 days
Cvetković et al (2016) [47]	9	0 diabetic and 9 healthy; 1 female and 8 males; and aged 24-36 years	14 days
Calbimonte et al (2017) [38]	Not described	External data set	Not described
Fraiwan et al (2017) [48]	Not described	Healthy; gender undefined; and age undefined	Not described
McLean et al (2017) [49]	22	22 diabetic and 0 healthy; gender undefined; and age undefined	Not described
Razjouyan et al (2017) [57]	25	25 diabetic and 0 healthy; gender undefined; aged 59.3 (SD 8.3) years	21 (SD 4) days
Reddy et al (2017) [39]	100	50 diabetic and 50 healthy; gender undefined; and aged 34 (SD 10) years (diabetic) and 41 (SD 13) years (healthy)	5 mnin
Turksoy et al (2017) [50]	26	26 diabetic and 0 healthy; 14 females and 12 males; and aged 24.2 (SD 5.41) years	6 days
Bartolic et al (2018) [40]	Not described	Health status undefined; gender undefined; and age undefined	Not described
Faccioli et al (2018) [45]	6	6 diabetic and 0 healthy; gender undefined; and age undefined	5 days
Groat et al (2018) [51]	12	12 diabetic and 0 healthy; 8 females and 4 males; and aged 48 (SD 13.4) years	30 days
McMillan et al (2018) [46]	1	1 diabetic and 0 healthy; 0 female and 1 male; and aged 68 years	1 day
Merickel et al (2018) [32]	36	20 diabetic and 16 healthy; gender undefined; and aged 21-59 years	28 days
Nguyen Gia et al (2019) [41]	4	0 diabetic and 4 healthy; gender undefined; aged 30 years	Not described
Rescio et al (2019) [33]	5	Health status undefined; gender undefined; and aged 47.2 (SD 12.3) years	Not described
Sarda et al (2019) [52]	46	46 diabetic and 0 healthy; 17 females and 29 males; and aged 35 (SD 12) years	140 days
Ramazi et al (2019) [34]	50	50 diabetic and 0 healthy; gender undefined; and aged 33-78 years	7 days
Garcia et al (2019) [35]	100	50 diabetic and 50 healthy; 58 females and 42 males; and aged 20-87 years	Not described
Sevil et al (2019) [42]	2	2 diabetic and 0 healthy; gender undefined; and age undefined	1 day
Zherebtsov et al (2019) [43]	55	18 diabetic and 37 healthy; gender undefined; and aged 53.2 (SD 11.4) years (diabetic), 19.6 (SD 0.6) years (16 healthy), and 53.2 (SD 11.4) years (21 healthy)	10 min
Rodriguez-Rodriguez et al (2019) [53]	25	25 diabetic and 0 healthy; 11 females and 14 males; and aged 18-56 years	14 days
Sanz et al (2019) [54]	6	6 diabetic and 0 healthy; gender undefined; and aged 36.7 (SD 8.9) years	Not described
Whelan et al (2019) [55]	45	0 diabetic and 45 healthy; 27 females and 18 males; and aged 56 (SD 9) years	42 days

Discussion

Principal Findings

The reviewed studies revealed the potential of mobile and wearable technologies in health areas. These technologies can significantly improve the management of conditions for both patients and clinicians for a variety of diseases. DM is not an exception, and growing attention has been paid to the use of these technologies in the recent years. Obtaining objective and continuous measurements is an important advantage of using this technology for patient monitoring. Data are sensed automatically by electronic sensors when the subject is interacting with the mobile or wearable devices both explicitly and implicitly, such as phone calls or step counts, respectively. These technologies most often enable the seamless collection of data, even when the patient is out of the clinic. This is a relevant feature to overcome the drawbacks of classical clinical trials in which subjects are required to stay in labs or clinics, set specific appointments, commute to the doctor's office, etc. This technology adds a level of objectivity in the monitoring of patients with DM and people in general with respect to traditional clinical questionnaires, which are more dependent on users' willingness and capacity to answer correctly. In addition, the patient may not remember everything accurately in between doctor visits and hospitalization times. In view of such limitations, the opportunities for the monitoring of DM-related parameters are unprecedented. Nonetheless, it is clear from this review that mobile and wearable technologies have been scarcely exploited for this purpose.

Several studies have not indicated the type of DM. In such studies, the authors refer to the condition simply as *diabetes* or *diabetes mellitus*. We assume that in those cases, the results of the research are generic and can be applied to any type of

diabetic. However, there was generally a lack of clinical specifications, perhaps because the majority of studies were published in technological journals and proceedings and, in part, because of the recency of this new field. In either case, such a level of specification is considered of utmost importance because of the differences among the types of diabetes and their decompensations, risk factors, age of onset, and treatment among several others. Complications of the disease may result in both types of DM, but few studies have focused on this subject. DPN was the only complication found in the reviewed studies and in almost all cases related to diabetic foot disease. Some important complications such as retinopathy; nephropathy; and other types of neuropathies, such as autonomic, focal, and proximal, were not considered in any of the selected studies. Few studies have elaborated on T2D, the type encompassing between 90% and 95% of the cases of DM in the world and the only preventable one.

Clinical trials were quite limited or even nonexistent in many of the reviewed studies. In fact, the majority of analyzed contributions had a predominant technological focus, prioritizing systems' performance or robustness over the impact or applicability of potential clinical outcomes. This explains the remarkably low scores achieved by most of the selected studies in the NOS quality assessment. Most of the reviewed cohort studies did not have a sufficiently representative cohort. Most often, a distinction between exposed and nonexposed cohorts was not clearly made, or even worse, no description of the derivation of the cohort was provided. This lack of detail was also observed for cases and controls. Comparability was also found to be quite limited, as exposed and nonexposed individuals, if any, were not matched in the design, and confounders were either missing or not adjusted for in the analysis. Although outcomes were assessed in a majority of studies, follow-ups were mostly nonexistent or no information was provided whatsoever. Therefore, one of the major weaknesses detected in the reviewed studies is the limited dedication to the clinical validation of the proposed technical solutions.

Mobile and wearable data can shed new light on behavioral and physiological aspects that are difficult to approach in a continuous and unobtrusive manner via standard clinical tests. However, ignoring clinical data is certainly a big mistake. Therefore, combining passive mobile data with clinical data, such as laboratory test results, drug information, or patient demographics, is key for a holistic understanding of the patient's current and future health status. Thus, it is recommended to perform more extensive clinical tests and validations involving the collection of new data sets. Existing data sets in this area show important flaws such as noninclusion of patients with DM, noninclusion of complementary clinical data, lack of gender diversity, or age variety. Moreover, public sharing of data sets is also considered essential to facilitate the replicability and reproducibility of the studies. Hence, data transparency and openness are encouraged, as in other similar disciplines.

None of the reviewed studies focused on the prevention of DM. This matter is especially important in the case of T2D, the only type of DM that can be prevented. Therefore, developing studies with outcomes that help to detect the disease in the early stages or even before it occurs can result in great progress. Approaching this subject from a holistic perspective could also be key for making new successful findings. This is closely related to the idea of using different data sources to generate more powerful medical models. Combining demographics, nutrition data, medication data, and passive sensor data among other heterogeneous data types can certainly help to realize more impactful and personalized solutions.

Activity recognition is one of the most important areas from which the monitoring of DM-related parameters could benefit. Thus, the research conducted in this new field may not only leverage the results from previous studies but also help in developing and testing new activity detection models. For example, improvements in the recognition of eating activities are needed to calculate food intake automatically.

Most of the studies had technological test phases, but in some cases, their quality was rather questionable. The description was often incomplete, lacked characteristics of the subjects, and did not mention the duration of the tests in several cases. These seem to be characteristics deemed in studies in the early stages. However, many of these studies stated that improvements would be made regarding this aspect in future research.

The studies analyzed in this review applied a variety of devices and sensors. Some case studies only used smartphones, others used only wearables, and others used a combination of them. This shows the ways in which these devices can be used to improve DM control and its complications. However, there is generally a poor description of the devices used in terms of their brand, model, manufacturer, main features, operating system, etc. This is an especially sensitive barrier for the replication of studies and development of follow-up research. Likewise, on some occasions, the sensors embedded in these devices were not explicitly described. Some studies included complementary devices such as a pressure platform, a glucometer, or an ambient sensor, not necessarily wearables, which helped to obtain more complete data and better characterize the patient's environment. Most often, all devices were merely used to collect data for creating ML models and to find an association among variables or diseases, but in very few cases, the proposed solution was implemented in a realistic use case with long-lasting clinical applications.

The predominant sensor was the ACC, possibly because it is one of the most common sensors available on both wearables and mobile devices. In addition, its applications are closely related to energy expenditure and activity recognition tasks, which are very useful in DM problems. Other sensors such as GPS, thermometer, microphone, and ambient light are less commonly used in the reviewed studies. This may be because some of these sensors, such as GPS and microphone, are considered to be more privacy-intrusive by users. Nonetheless, these sensors were shown to be helpful as complements to other sensors for the monitoring of DM-related parameters. Furthermore, it is worth noting that none of the studies used commercial smart insoles but one prototype, especially given the fact that most important complications of DM translate into foot issues.

Diabetes technology has grown in the recent years, with CGM being at the forefront of the devices used. CGM is primarily used to monitor patients with T1D, with increasing use for patients with T2D. However, the use of CGM does not replace the traditional finger-stick test because patients still need to do a meter reading for accuracy, and in most cases, insurance companies do not cover the use of CGM. The fact that mobile and wearable technologies are at the reach of a majority of the population makes the reviewed solutions particularly cost-effective.

The power consumption of the hardware available in both mobiles and wearables has decreased with the latest advances. However, there is still much to be done to reduce it even further. The 5G technology promises to do so by lowering network energy use by almost 90% and increasing battery life, especially for low-power devices [58]. Allowing patients to use wearable devices for prolonged periods without recharging could help foster their use in the diabetic community. The 5G technology will also increase the connection speed and bandwidth in a unit area and will have a very low latency. This will provide higher real-time monitoring capabilities, an increase in the amount of data sent by time unit, and the possibility of having more wearable devices connected to the cloud without requiring the use of smartphones as gateways. All this provides more self-sufficiency to these devices.

Data collected using mobile and wearable devices for continuous monitoring can be mined using AI techniques such as ML. As shown in this review, some authors used ML to extract information from the collected data, where large and heterogeneous data sets generally improve the performance of these techniques. Data from mobile and wearable sensors can, in principle, be used in combination with conventional clinical data to develop more relevant knowledge outcomes. The time and effort required to collect a data set that can be used to apply ML techniques is reduced by the use of mobile and wearable devices. Classical approaches are generally constrained by the number of samples or data points, as these are measured during clinic appointments, thus leading to lengthier collection phases.

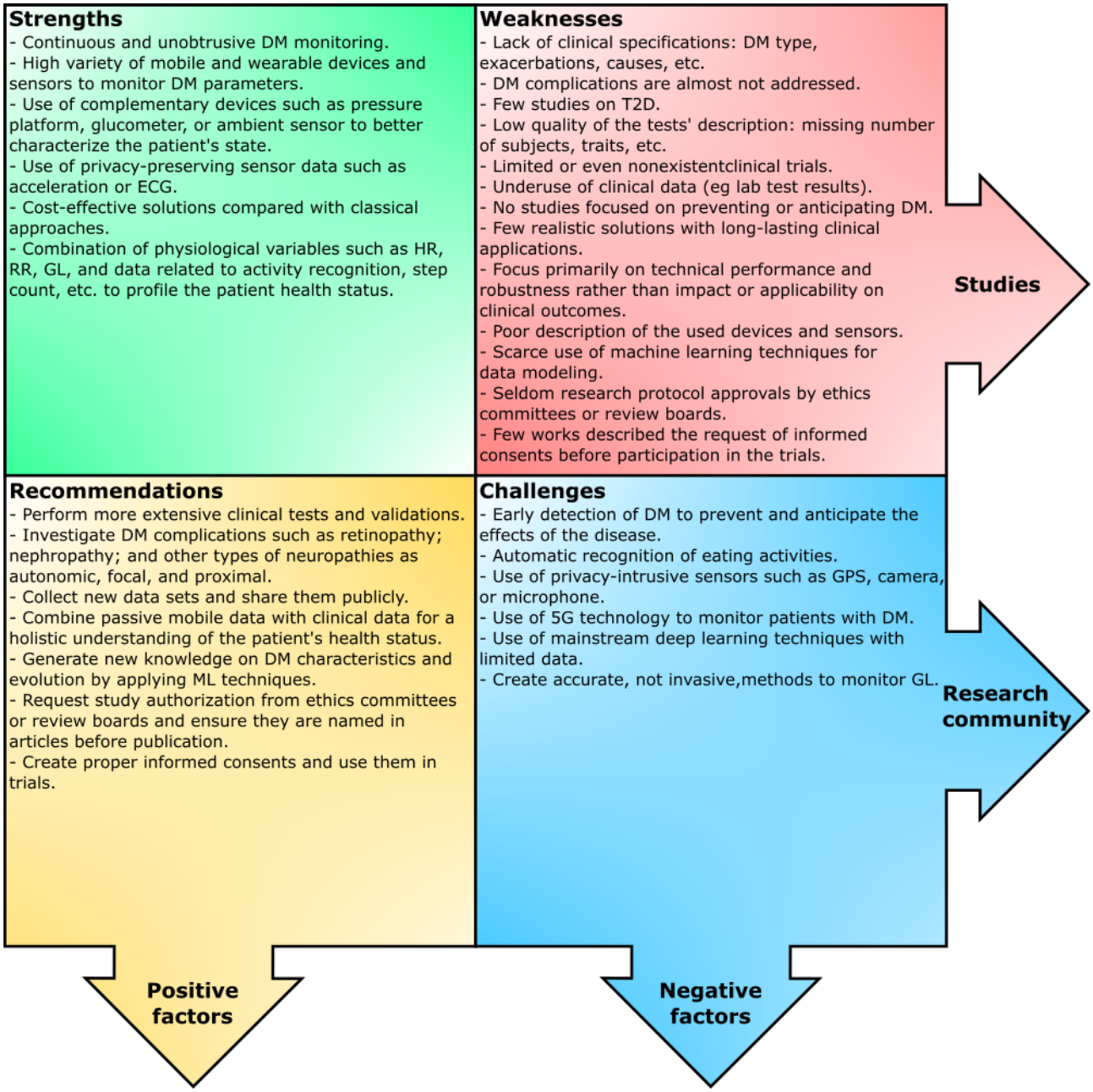
The most common ML techniques used in these studies were decision trees and support vector machines, whereas other popular algorithms such as k-nearest neighbors, artificial neural networks, ensemble methods, and deep learning techniques are used less frequently. These techniques, especially when it comes to deep learning, can significantly boost the performance of the results, normally at the expense of having large data sets, a condition that is normally attained when using passive sensing. In general, experimentation with ML algorithms was performed using a small number of methods, whereas the use of a large variety of these techniques normally leads to higher robustness. Very few studies have used complementary clinical data in addition to sensor data, resulting in better models and more relevant outcomes.

Many manuscripts did not mention the endorsement of their studies from an ethics committee or review board. This is especially important because in several cases, people were used to test the proposed solutions. It is even more important to highlight that in very few occasions, the studies described that informed consent was requested from the participants in the trials. This occurs even when sensitive information is recorded from users during monitoring, such as location, video, or call logs. A reason for this could be linked to the rather emerging nature of this field or the lack of realistic clinical studies around which technical solutions have been developed.

Privacy and security issues were weak aspects of the reviewed studies. Researchers should devote more attention to both realizing and explaining proper procedures to ensure that security and privacy are properly addressed in clinical studies. Otherwise, the quality and applicability of the results are compromised. An effort must be made to put into practice the protocols in the trial involving the ethics committee or review boards in the authorization of the studies. Creating proper informed consents forms and using them in the trials should be a major concern for research in this area, especially in lieu of regulations such as the European General Data Protection Regulation. Similarly, information on data management plans can provide further details on how the research has been undertaken.

A summary of the principal findings described previously is provided in Figure 2. The diagram shows the strengths and weaknesses of the reviewed studies on the monitoring of DM-related parameters as well as the challenges and recommendations for the research community. Strengths are the aspects that these studies have performed well on and could be reproduced in future investigations. Weaknesses are matters that went wrong in these studies and could be improved in future research. Challenges are the elements that the scientific community needs to address successfully to boost the investigation of this topic. Recommendations are the suggestions for the research community working on the monitoring of DM-related parameters.

Figure 2. Summary of the principal findings of the reviewed manuscripts. DM: diabetes mellitus; ECG: electrocardiogram; GL: glucose level; HR: heart rate; ML: machine learning; RR: respiration rate; T2D: type 2 diabetes.



Limitations

The monitoring of DM-related parameters using mobile and wearable technology is an emerging field of study. As for any other review, despite having listed a wide variety of terms referring to sensors, wearable devices, and smartphones, new keywords emerge quite often in this rather dynamic technological area, which may have left out some interesting studies from our analysis. Although the search areas of this systematic review (computer science and engineering) are quite large, it is also possible that some relevant studies indexed in other related categories may have been filtered out. We conducted a preliminary check for other domains such as endocrinology metabolism, general internal medicine, or health care sciences services, and we did not find relevant studies that would meet the defined criteria. Other sources of digital data, such as social network interactions, have not been considered in this study, as they can be realized via different technologies besides mobiles and wearables. Nevertheless, it could be interesting to explore the potential of these interactions to explain some relevant behaviors of the patient, such as their mood. The lack of details in some studies also made it difficult to judge whether the authors were using commercial devices or their own prototype. Thus, it is possible that some relevant studies were excluded, although this is in line with the PRISMA guidelines followed in this systematic review.

Conclusions

As demonstrated in this systematic review, the field of mobile and wearable monitoring of DM-related parameters shows early promise, despite its recent development. Several actors may benefit at the maturation of this field: (1) patients with DM, who may have a better quality of life while improving the management and self-control of the disease or its complications

in a continuous, passive, and unobtrusive way; (2) health care professionals and institutions, who may develop the ability to provide medical care and information in a portable and affordable way; and (3) researchers, who may have access to a large and varied amount of data sets to extract relevant information. The aforementioned 3 actors may work in synergy, which motivates a greater and faster evolution of the field. However, some gaps remain to accomplish this view, such as the creation or modification of relevant sensors to be less privacy-intrusive; decreasing the devices' power consumption; using the advantages of the 5G technology; and, perhaps the most important one, combining passive mobile data with clinical data for a holistic understanding of the patient's health status. Accomplishing these challenges requires interdisciplinary teams' collaboration and the appropriate funding of governments and institutions to design and develop the required technologies for sensing the data, designing new and better processing techniques, and creating realistic solutions with long-lasting clinical applications.

Acknowledgments

This study was funded by the University of Granada within the framework of the Development Cooperation Fund. The study was also partially funded by the Spanish Project Advanced Computing Architectures and Machine Learning-Based Solutions for Complex Problems in Bioinformatics, Biotechnology, and Biomedicine (RTI2018-101674-B-I00).

Authors' Contributions

CR and OB conceptualized the work. OB and CV defined the review methodology. CR and OB conducted screening at all levels and data extraction with the help of CV. CR, CV, and OB developed the analysis protocol and conducted the systematic analysis. CR and JRR conducted the quality assessment. CR, CV, MM, JRR and OB contributed to manuscript drafting, review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Newcastle-Ottawa Scale quality assessment of the selected studies.
[PDF File (Adobe PDF File), 74 KB - mhealth_v9i6e25138_app1.pdf]

Multimedia Appendix 2
Reviewed manuscripts by the year of publication (up to July 28, 2020).
[PNG File , 32 KB - mhealth_v9i6e25138_app2.png]

Multimedia Appendix 3
Sensor use for each study.
[PDF File (Adobe PDF File), 274 KB - mhealth_v9i6e25138_app3.pdf]

References

1. Baena-Díez JM, Peñafiel J, Subirana I, Ramos R, Elosua R, Marín-Ibañez A, FRESCO Investigators. Risk of cause-specific death in individuals with diabetes: a competing risks analysis. Diabetes Care 2016 Nov;39(11):1987-1995. [doi: 10.2337/dc16-0614] [Medline: 27493134]
2. IDF Diabetes Atlas: Ninth Edition 2019. Belgium: International Diabetes Federation; 2019:1-176.
3. Zimmet PZ, Magliano DJ, Herman WH, Shaw JE. Diabetes: a 21st century challenge. Lancet Diabetes Endocrinol 2014 Jan;2(1):56-64. [doi: 10.1016/S2213-8587(13)70112-8] [Medline: 24622669]
4. IDF Diabetes Atlas 8th Ed 2017. Belgium: International Diabetes Federation; 2017:1-150.
5. Classification of Diabetes Mellitus 2019. Geneva: World Health Organization; 2019:1-40.
6. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2008 Jan;31 Suppl 1:55-60. [doi: 10.2337/dc08-S055] [Medline: 18165338]
7. World Health Organization. Mortality and Burden of Disease Attributable to Selected Major Risks. Geneva: World Health Organization; 2009:1-62.
8. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. J Med Internet Res 2018 May 30;20(5):e10775 [FREE Full text] [doi: 10.2196/10775] [Medline: 29848472]
9. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med 1998 Jul;15(7):539-553. [doi: 10.1002/(SICI)1096-9136(199807)15:7<539::AID-DIA668>3.0.CO;2-S] [Medline: 9686693]
10. Diabetes. World Health Organization. URL: https://www.who.int/diabetes/action_online/basics/en/index3.html [accessed 2020-07-24]

11. Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. N Engl J Med 2001 Sep 13;345(11):790-797. [doi: 10.1056/NEJMoa010492] [Medline: 11556298]
12. d'Emden MC, Shaw JE, Jones GR, Cheung NW. Guidance concerning the use of glycated haemoglobin (HbA1c) for the diagnosis of diabetes mellitus. Med J Aust 2015 Jul 20;203(2):89-90. [doi: 10.5694/mja15.00041] [Medline: 26175248]
13. O'Connor PJ, Bodkin NL, Fradkin J, Glasgow RE, Greenfield S, Gregg E, et al. Diabetes performance measures: current status and future directions. Diabetes Care 2011 Jul;34(7):1651-1659 [FREE Full text] [doi: 10.2337/dc11-0735] [Medline: 21709298]
14. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
15. American Diabetes Association. 7. Diabetes Technology: standards of medical care in diabetes-2020. Diabetes Care 2020 Jan;43(Suppl 1):77-88. [doi: 10.2337/dc20-S007] [Medline: 31862750]
16. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
17. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-06-24]
18. Global attitudes and trends. Pew Research Center. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2019-06-24]
19. Wearables - worldwide. Statista. URL: https://www.statista.com/outlook/319/100/wearables/worldwide [accessed 2019-07-25]
20. Godfrey A, Hetherington V, Shum H, Bonato P, Lovell NH, Stuart S. From A to Z: wearable technology explained. Maturitas 2018 Jul;113:40-47 [FREE Full text] [doi: 10.1016/j.maturitas.2018.04.012] [Medline: 29903647]
21. Richter F. The global wearables market is all about the wrist. Statista. 2018. URL: https://www.statista.com/chart/3370/wearable-device-forecast/ [accessed 2019-07-25]
22. Felix IR, Castro LA, Rodriguez L, Banos O. Mobile sensing for behavioral research: a component-based approach for rapid deployment of sensing campaigns. Int J Distrib Sens N 2019 Sep 06;15(9). [doi: 10.1177/1550147719874186]
23. Banos O, Villalonga C, Bang J, Hur T, Kang D, Park S, et al. Human behavior analysis by means of multimodal context mining. Sensors (Basel) 2016 Aug 10;16(8):1264 [FREE Full text] [doi: 10.3390/s16081264] [Medline: 27517928]
24. Lee U, Han K, Cho H, Chung K, Hong H, Lee S, et al. Intelligent positive computing with mobile, wearable, and IoT devices: literature review and research directions. Ad Hoc Networks 2019 Feb;83:8-24. [doi: 10.1016/j.adhoc.2018.08.021]
25. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. J Diabetes 2018 Jan;10(1):10-17. [doi: 10.1111/1753-0407.12606] [Medline: 28872765]
26. Yu K, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018 Oct;2(10):719-731 [FREE Full text] [doi: 10.1038/s41551-018-0305-z] [Medline: 31015651]
27. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243 [FREE Full text] [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
28. Rumbold JM, O'Kane M, Philip N, Pierscionek BK. Big Data and diabetes: the applications of Big Data for diabetes care now and in the future. Diabet Med 2020 Feb;37(2):187-193. [doi: 10.1111/dme.14044] [Medline: 31148227]
29. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
30. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2021-05-11]
31. Taskiran NO, editor. Advances in Media, Entertainment, and the Arts (AMEA) book series. In: Handbook of Research on Multidisciplinary Approaches to Literacy in the Digital Age. Barcelona: IGI Global; 2020:1-405.
32. Merickel J, High R, Smith L, Wichman C, Frankel E, Smits K, et al. Driving Safety and Real-Time Glucose Monitoring in Insulin-Dependent Diabetes. Int J Automot Eng 2019;10(1):34-40. [doi: 10.20485/jsaeijae.10.1_34]
33. Rescio G, Leone A, Francioso L, Siciliano P. Smart insole for diabetic foot monitoring. In: Lecture Notes in Electrical Engineering. Switzerland: Springer; 2019:571-577.
34. Ramazi R, Perndorfer C, Soriano E, Laurenceau JP, Beheshti R. Multi-modal predictive models of diabetes progression. In: Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics. 2019 Presented at: BCB '19: 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics; September, 2019; Niagara Falls NY USA p. 253-258. [doi: 10.1145/3307339.3342177]
35. Garcia CA. Non-invasive diabetes detection using facial texture features captured in a less restrictive environment. Int J Recent Technol Eng 2019 Nov 2;8(2S11):3469-3473. [doi: 10.35940/ijrte.b1584.0982s1119]
36. Luštrek M, Cvetković B, Janko V. Monitoring patients' lifestyle with a smartphone and other devices placed freely on the body. In: Ambient Intelligence. Switzerland: Springer; 2014:96-99.
37. Luštrek M, Cvetkovic B, Mirchevska V, Kafalı O, Romero A, Stathis K. Recognising lifestyle activities of diabetic patients with a smartphone. European Union Digital Library 2015. [doi: 10.4108/icst.pervasivehealth.2015.259118]

38. Calbimonte J, Ranvier J, Dubosson F, Aberer K. Semantic representation and processing of hypoglycemic events derived from wearable sensor data. J Ambient Intell Smart Environ 2017 Jan 16;9(1):97-109. [doi: 10.3233/ais-160420]
39. Reddy VR, Choudhury AD, Deshpande P, Jayaraman S, Thokala NK, Kaliaperumal V. Dmsense: a non-invasive diabetes mellitus classification system using photoplethysmogram signal. In: Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). 2017 Presented at: IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); March 13-17, 2017; Kona, HI, USA. [doi: 10.1109/percomw.2017.7917526]
40. Bartolic L, Zorić B, Martinović G. E-Gluko: a ubiquitous system for health status monitoring and tracking in diabetes patients. In: Proceedings of the International Conference on Smart Systems and Technologies (SST). 2018 Presented at: International Conference on Smart Systems and Technologies (SST); Oct. 10-12, 2018; Osijek, Croatia. [doi: 10.1109/sst.2018.8564570]
41. Gia TN, Dhaou IB, Ali M, Rahmani AM, Westerlund T, Liljeberg P, et al. Energy efficient fog-assisted IoT system for monitoring diabetic patients with cardiovascular disease. Future Gener Comput Syst 2019 Apr;93:198-211. [doi: 10.1016/j.future.2018.10.029]
42. Sevil M, Rashid M, Hajizadeh I, Maloney Z, Samadi S, Askari MR, et al. Assessing the effects of stress response on glucose variations. In: Proceedings of the 16th IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN). 2019 Presented at: 16th IEEE International Conference on Wearable and Implantable Body Sensor Networks; May 19-22, 2019; Chicago, IL, USA. [doi: 10.1109/bsn.2019.8771094]
43. Zherebtsov EA, Zharkikh EV, Kozlov IO, Loktionova YI, Zherebtsova AI, Rafailov IE, et al. Wearable sensor system for multipoint measurements of blood perfusion: pilot studies in patients with diabetes mellitus. In: Medical Laser Applications and Laser-Tissue Interactions. Bellingham: SPIE Proceedings (Optical Society of America, 2019); 2019.
44. Najafi B, Horn D, Marclay S, Crews RT, Wu S, Wrobel JS. Assessing postural control and postural control strategy in diabetes patients using innovative and wearable technology. J Diabetes Sci Technol 2010 Jul 01;4(4):780-791 [FREE Full text] [doi: 10.1177/193229681000400403] [Medline: 20663438]
45. Faccioli S, Ozaslan B, Garcia-Tirado JF, Breton M, Favero SD. Black-box model identification of physical activity in type-1 diabetes patients. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018 Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18-21, 2018; Honolulu, HI, USA. [doi: 10.1109/embc.2018.8513378]
46. McMillan KA, Kirk A, Hewitt A, MacRury S, Lennon M. Methods for combining continuously measured glucose and activity data in people with type 2 diabetes: challenges and solutions. J Rehabil Assist Technol Eng 2018;5 [FREE Full text] [doi: 10.1177/2055668318782805] [Medline: 31191946]
47. Cvetković B, Janko V, Romero AE, Kafalı O, Stathis K, Luštrek M. Activity recognition for diabetic patients using a smartphone. J Med Syst 2016 Dec;40(12):256. [doi: 10.1007/s10916-016-0598-y] [Medline: 27722975]
48. Fraiwan L, AlKhodari M, Ninan J, Mustafa B, Saleh A, Ghazal M. Diabetic foot ulcer mobile detection system using smart phone thermal camera: a feasibility study. Biomed Eng Online 2017 Oct 03;16(1):117 [FREE Full text] [doi: 10.1186/s12938-017-0408-x] [Medline: 28974212]
49. McLean A, Osgood N, Newstead-Angel J, Stanley K, Knowles D, van der Kamp W, et al. Building research capacity: results of a feasibility study using a novel mhealth epidemiological data collection system within a gestational diabetes population. Stud Health Technol Inform 2017;234:228-232. [Medline: 28186046]
50. Turksoy K, Monforti C, Park M, Griffith G, Quinn L, Cinar A. Use of wearable sensors and biometric variables in an artificial pancreas system. Sensors (Basel) 2017 Mar 07;17(3):532 [FREE Full text] [doi: 10.3390/s17030532] [Medline: 28272368]
51. Groat D, Kwon HJ, Grando MA, Cook CB, Thompson B. Comparing real-time self-tracking and device-recorded exercise data in subjects with type 1 diabetes. Appl Clin Inform 2018 Oct;9(4):919-926 [FREE Full text] [doi: 10.1055/s-0038-1676458] [Medline: 30586673]
52. Sarda A, Munuswamy S, Sarda S, Subramanian V. Using passive smartphone sensing for improved risk stratification of patients with depression and diabetes: cross-sectional observational study. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11041 [FREE Full text] [doi: 10.2196/11041] [Medline: 30694197]
53. Rodríguez-Rodríguez I, Chatzigiannakis I, Rodríguez JV, Maranghi M, Gentili M, Zamora-Izquierdo M. Utility of big data in predicting short-term blood glucose levels in type 1 diabetes mellitus through machine learning techniques. Sensors (Basel) 2019 Oct 16;19(20):4482 [FREE Full text] [doi: 10.3390/s19204482] [Medline: 31623111]
54. Sanz AJ, Díez JL, Giménez M, Bondia J. Enhanced accuracy of continuous glucose monitoring during exercise through physical activity tracking integration. Sensors (Basel) 2019 Aug 30;19(17):3757 [FREE Full text] [doi: 10.3390/s19173757] [Medline: 31480343]
55. Whelan ME, Orme MW, Kingsnorth AP, Sherar LB, Denton FL, Esliger DW. Examining the use of glucose and physical activity self-monitoring technologies in individuals at moderate to high risk of developing type 2 diabetes: randomized trial. JMIR Mhealth Uhealth 2019 Oct 28;7(10):e14195 [FREE Full text] [doi: 10.2196/14195] [Medline: 31661077]

56. Grewal GS, Bharara M, Menzies R, Talal TK, Armstrong D, Najafi B. Diabetic peripheral neuropathy and gait: does footwear modify this association? J Diabetes Sci Technol 2013 Sep 01;7(5):1138-1146 [FREE Full text] [doi: 10.1177/193229681300700506] [Medline: 24124939]
57. Razjouyan J, Grewal GS, Talal TK, Armstrong DG, Mills JL, Najafi B. Does physiological stress slow down wound healing in patients with diabetes? J Diabetes Sci Technol 2017 Jul;11(4):685-692 [FREE Full text] [doi: 10.1177/1932296817705397] [Medline: 28436270]
58. Agiwal M, Saxena N, Roy A. Ten commandments of emerging 5G networks. Wireless Pers Commun 2017 Oct 30;98(3):2591-2521. [doi: 10.1007/s11277-017-4991-8]

Abbreviations

ACC: accelerometer
AI: artificial intelligence
CGM: continuous glucose monitor
DFU: diabetic foot ulcer
DM: diabetes mellitus
DPN: diabetic peripheral neuropathy
ECG: electrocardiogram
FGM: flash glucose monitor
GD: gestational diabetes
GL: glucose level
GM: glucose monitor
HR: heart rate
ML: machine learning
NOS: Newcastle-Ottawa Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RR: respiration rate
T1D: type 1 diabetes
T2D: type 2 diabetes
WoS: Web of Science

Edited by L Buis; submitted 19.10.20; peer-reviewed by Z Ren, E Drokow; comments to author 26.10.20; revised version received 27.11.20; accepted 11.03.21; published 03.06.21.

Please cite as:
Rodriguez-León C, Villalonga C, Munoz-Torres M, Ruiz JR, Banos O
Mobile and Wearable Technology for the Monitoring of Diabetes-Related Parameters: Systematic Review
JMIR Mhealth Uhealth 2021;9(6):e25138
URL: https://mhealth.jmir.org/2021/6/e25138
doi:10.2196/25138
PMID:34081010

Review

Best Practice Guidance for Digital Contact Tracing Apps: A Cross-disciplinary Review of the Literature

James O'Connell[1], MB BCh BAO, MSc; Manzar Abbas[1], BSc, MSc; Sarah Beecham[1], PhD; Jim Buckley[1], PhD; Muslim Chochlov[1], PhD; Brian Fitzgerald[1], PhD; Liam Glynn[2], MD; Kevin Johnson[3], BEng, MEng, PhD; John Laffey[4,5], BSc, MB, MA, MD; Bairbre McNicholas[4,5], BSc, MB BCh, PhD; Bashar Nuseibeh[1,6], PhD, MAE, MRIA; Michael O'Callaghan[2], BM, BS, BEng, MSc; Ian O'Keeffe[1], BA, BAI, MSc, PhD; Abdul Razzaq[1], PhD; Kaavya Rekanar[1], BTech, MSc; Ita Richardson[1], BSc, MSc, PhD; Andrew Simpkin[7], PhD; Cristiano Storni[1], PhD; Damyanka Tsvyatkova[1], PhD; Jane Walsh[8], PhD; Thomas Welsh[1], PhD; Derek O'Keeffe[1,4,5], BEng, MB BCh BAO, MEng, MD, PhD

[1]Lero, Science Foundation Ireland Research Centre for Software, University of Limerick, Limerick, Ireland

[2]School of Medicine, University of Limerick, Limerick, Ireland

[3]Department of Nursing and Midwifery, University of Limerick, Limerick, Ireland

[4]School of Medicine, National University of Ireland Galway, Galway, Ireland

[5]University Hospital Galway, Saolta, Health Services Executive, Galway, Ireland

[6]School of Computing and Communications, The Open University, Milton Keynes, United Kingdom

[7]School of Mathematics, Statistics and Applied Mathematics, National University of Ireland Galway, Galway, Ireland

[8]School of Psychology, National University of Ireland Galway, Galway, Ireland

Corresponding Author:
Derek O'Keeffe, BEng, MB BCh BAO, MEng, MD, PhD
School of Medicine
National University of Ireland Galway
Clinical Sciences Institute
Costello Road
Galway, H91 V4AY
Ireland
Phone: 353 (0)91 495960
Email: derek.okeeffe@nuigalway.ie

Abstract

Background: Digital contact tracing apps have the potential to augment contact tracing systems and disrupt COVID-19 transmission by rapidly identifying secondary cases prior to the onset of infectiousness and linking them into a system of quarantine, testing, and health care worker case management. The international experience of digital contact tracing apps during the COVID-19 pandemic demonstrates how challenging their design and deployment are.

Objective: This study aims to derive and summarize best practice guidance for the design of the ideal digital contact tracing app.

Methods: A collaborative cross-disciplinary approach was used to derive best practice guidance for designing the ideal digital contact tracing app. A search of the indexed and gray literature was conducted to identify articles describing or evaluating digital contact tracing apps. MEDLINE was searched using a combination of free-text terms and Medical Subject Headings search terms. Gray literature sources searched were the World Health Organization Institutional Repository for Information Sharing, the European Centre for Disease Prevention and Control publications library, and Google, including the websites of many health protection authorities. Articles that were acceptable for inclusion in this evidence synthesis were peer-reviewed publications, cohort studies, randomized trials, modeling studies, technical reports, white papers, and media reports related to digital contact tracing.

Results: Ethical, user experience, privacy and data protection, technical, clinical and societal, and evaluation considerations were identified from the literature. The ideal digital contact tracing app should be voluntary and should be equitably available and accessible. User engagement could be enhanced by small financial incentives, enabling users to tailor aspects of the app to

their particular needs and integrating digital contact tracing apps into the wider public health information campaign. Adherence to the principles of good data protection and privacy by design is important to convince target populations to download and use digital contact tracing apps. Bluetooth Low Energy is recommended for a digital contact tracing app's contact event detection, but combining it with ultrasound technology may improve a digital contact tracing app's accuracy. A decentralized privacy-preserving protocol should be followed to enable digital contact tracing app users to exchange and record temporary contact numbers during contact events. The ideal digital contact tracing app should define and risk-stratify contact events according to proximity, duration of contact, and the infectiousness of the case at the time of contact. Evaluating digital contact tracing apps requires data to quantify app downloads, use among COVID-19 cases, successful contact alert generation, contact alert receivers, contact alert receivers that adhere to quarantine and testing recommendations, and the number of contact alert receivers who subsequently are tested positive for COVID-19. The outcomes of digital contact tracing apps' evaluations should be openly reported to allow for the wider public to review the evaluation of the app.

Conclusions: In conclusion, key considerations and best practice guidance for the design of the ideal digital contact tracing app were derived from the literature.

(JMIR Mhealth Uhealth 2021;9(6):e27753) doi:10.2196/27753

KEYWORDS

digital contact tracing; automated contact tracing; COVID-19; SARS-CoV-2; mHealth; mobile app; app; tracing; monitoring; surveillance; review; best practice; design

Introduction

Background

COVID-19 was declared a global pandemic by the World Health Organization (WHO) on March 11, 2020 [1]. COVID-19 has caused the death of over 3.4 million people worldwide as of May 26, 2021 [2]. Most people (as high as 80%) infected by COVID-19 will have no symptoms or mild-to-moderate symptoms [3,4]. Severe illness and death due to COVID-19 are more likely to occur with increasing age and comorbidities such as chronic heart or lung disease [3]. The WHO advise on four key actions to contain COVID-19: social distancing, rapid testing of those with symptoms, tracing of case contacts, and the isolation of suspected and confirmed cases [5]. People with COVID-19 are thought to be most infectious to others within the 2 days preceding symptom onset [3,6]. Onward transmission from individuals in the presymptomatic phase of infection is considered to be enough to sustain the pandemic even if isolation of symptomatic cases occurs [6]. This can be mitigated by stringent social distancing measures, but these come with considerable socioeconomic costs [7].

An effective “test and trace” system is key if the most restrictive social distancing measures such as national “stay at home” orders are to be avoided [8]. Manual contact tracing requires significant human and logistical resources, and its effectiveness depends on the availability and proficiency of contact tracing staff [8,9]. In addition, humans are fallible and prone to recall bias, meaning that not all contacts may be identified reliably in retrospect. It is also not possible in many situations to identify contacts unfamiliar to the case. For contact tracing purposes, the infectious period is considered to be up to 10 days after symptom onset and to begin from 2 days before symptom onset or if the person is asymptomatic from 2 days before testing [10,11]. The incubation period for COVID-19 can be up to 14 days (and longer in 5% of cases) [12], meaning that not all contacts are captured by this definition. COVID-19 has a serial interval as short as 3.2 days; therefore, contact tracing and quarantine of contacts must be rapid to disrupt transmission chains [13].

A digital contact tracing app (DCTA) is an app that can detect and trace other app-carrying individuals who have had contact with one another that would risk COVID-19 transmission if one were to be infected. Early in the pandemic, DCTAs were seen as a potentially innovative solution to contain COVID-19 by augmenting the effectiveness of manual contact tracing [14]. DCTAs could disrupt transmission chains by rapidly identifying secondary cases prior to the onset of infectiousness and linking them into a system of quarantine, testing, and health care worker case management [8,15-17]. A COVID-19 modeling study from the United Kingdom estimated that if a DCTA were used by 56% of the population, then the reproductive value of the virus could be reduced below 1.0, controlling the disease [18]. By October 13, 2020, there were 120 DCTAs in 71 countries [19], and within months of digital contact tracing use, some key challenges became evident.

Global Digital Contact Tracing App Deployment

On March 20, 2020, Singapore became the first country in the world to launch a national DCTA, TraceTogether [20]. TraceTogether was downloaded by over 1.1 million users within a month of launching, despite having technical limitations [21]. It required Apple iPhone users to have the app open in the foreground and caused significant battery drain [21]. In South Korea, an extensive electronic surveillance system was used. GPS-enabled location tracking, closed-circuit television recordings, and credit card transactions were used to aid contact tracing [22]. How these data were used by health authorities to warn others of potential exposure to COVID-19 may have breached the privacy of those infected and contributed to a growth in social stigma associated with the disease [23]. In Israel, a network-based mass surveillance system using mobile phone GPS technology was launched on March 16 to identify COVID-19 case contacts [24,25]. Authorities in Israel reported that, after 1 month of surveillance, 36.8% of COVID-19 cases notified were identified using the surveillance system [25], although the system did have a false-positive detection rate of

5% [25]. Significant privacy concerns were raised by opponents of Israel's surveillance system, but ultimately, the supreme court ruled in favor of its use provided it was supported by primary legislation [24-26]. In Norway, a DCTA that used Bluetooth Low Energy (LE) and GPS location tracking was launched on April 16 and was downloaded 1.6 million times [27]. However, the National Institute of Public Health was forced to abandon the app after data protection authorities deemed there was no evidence of its effectiveness to justify location data collection [27]. Norway remained without a DCTA for several months thereafter [27]. Qatar mandated the use of its "Ehteraz" DCTA on May 22, but subsequently, it was discovered that it left the health status and location data of over 1 million users vulnerable to cyberattacks [28]. On July 7, the Republic of Ireland (ROI) launched a national DCTA called COVID Tracker, which is actively used by 1.3 million people, 34% of those older than 16 years nationally [29]. However, a service update in early August caused rapid battery depletion and heat issues for some users [30,31]. This was the primary cause of negative feedback for COVID Tracker [32]. In the 5-day period after this update, 152,656 uninstalls were registered with 29,049 returning users recorded [33]. As of May 28th 2021, eleven months after its release, COVID Tracker has been used by 15,742 people with COVID-19 to send contact alerts to 24,436 users [29]. In the United Kingdom, the National Health Service (NHS) COVID-19 app, launched on September 24, had been erroneously notifying users they were close contacts but provided no further instructions [34,35]. The international experience of DCTA use during the COVID-19 pandemic demonstrates how challenging their design and deployment are. This formed the basis of this literature review, which aims to derive and summarize best practice guidance for the design of the ideal DCTA (IDCTA).

Methods

A collaborative cross-disciplinary approach (Multimedia Appendix 1 [32,36-39]) was used to derive best practice guidance for designing the IDCTA. The cross-disciplinary team included specialists from computer science, engineering, clinical medicine, medical technology, and psychology. A scoping review to identify considerations described in the emerging literature on DCTAs was conducted (Multimedia Appendix 2 [8,14,15,36,40-42,44]). After the key considerations were identified and agreed upon by the cross-disciplinary team, a detailed evidence synthesis for each was constructed by author JOC and refined through a review and feedback cycle involving a subgroup of the cross-disciplinary team. The cycle of review and feedback was repeated until there was cross-disciplinary agreement that all feedback had been adequately addressed. The product of this process was then presented to the wider cross-disciplinary team for further discussion, from which best practice guidance for the design of the IDCTA was derived through a review and feedback cycle.

To construct the evidence synthesis, a literature search was conducted using Ovid MEDLINE and Epub Ahead of Print, In-Process, and Other Non-Indexed Citations, Daily and Versions. Free-text terms and Medical Subject Headings search terms were used (Multimedia Appendix 3). The WHO Institutional Repository for Information Sharing [45] and the European Centre for Disease Prevention and Control (ECDC) publications library [46] were searched. The gray literature search (Multimedia Appendix 4) included manually searching the websites of DCTAs, media sites, and health protection authorities including but not limited to the Centers for Disease Control and Prevention (United States of America), Public Health England, Health Protection Surveillance Centre (ROI), Robert Koch Institute (Germany), and the Norwegian Institute of Public Health. Included articles had to describe aspects of developing or deploying a DCTA. Peer-review publications, modeling studies, cohort studies, randomized trials, technical reports, white papers, and media reports were eligible for inclusion. The references of included articles were also searched to identify other eligible literature. Both English and non-English articles were included, and Google Translate was used to translate non-English articles.

Results

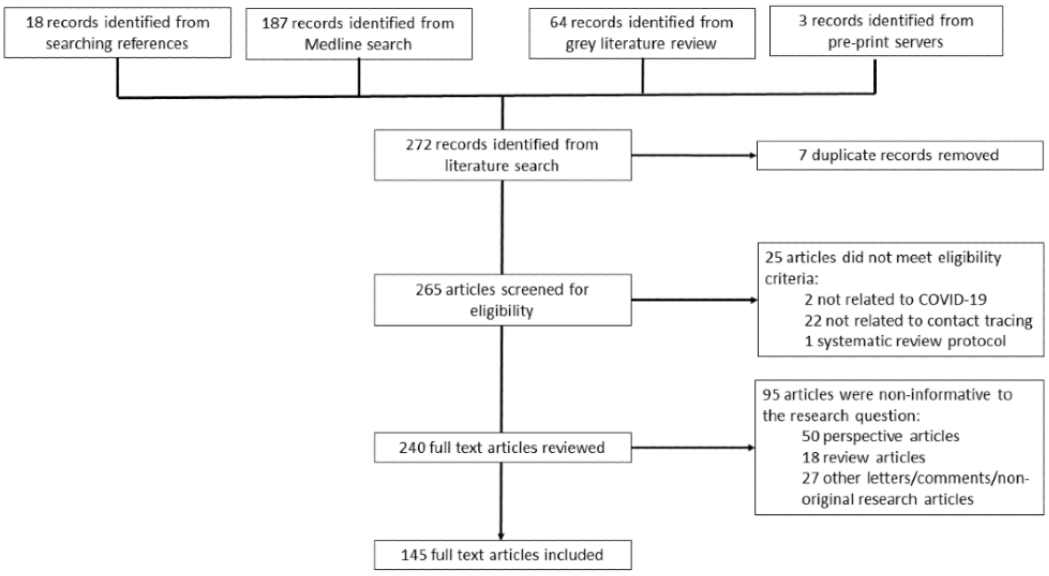
Key Considerations

From the scoping review (Multimedia Appendix 2), the cross-disciplinary team identified and agreed upon six key considerations for best practice guidance when designing the IDCTA: (1) ethical considerations, (2) user experience considerations, (3) privacy and data protection considerations, (4) technical considerations, (5) clinical and societal considerations, and (6) evaluation considerations.

The outcome of the literature search shown in Figure 1 and Multimedia Appendix 5 [14-17,22,23,25,26,32,34,37,38,40-42,44,47-173] contains a description of the included studies and their source (indexed literature search, gray literature, references search).

For each consideration, best practice guidance for the design of the IDCTA as derived from the literature by the cross-disciplinary group is summarized.

Figure 1. Literature search flow diagram.



Ethical Considerations

On December 10, 1948, the United Nations General Assembly adopted the Universal Declaration of Human Rights, which included the right to health [174]. Digital contact tracing can be viewed as states using their available resources to protect people's right to health. However, it may also be viewed as states interfering with other human rights enshrined in the Declaration, such as the right to human dignity; nondiscrimination; equality; privacy; access to information; and the freedoms of association, assembly, and movement [47,174]. Such interference to protect health may be ethical if it is adherent to criteria defined by the Siracusa Principles in 1985 [175]. These criteria are that interferences should have a legal basis, further a legitimate objective of common interest, not disrupt democratic processes, and not be deployed in an arbitrary or discriminatory way [175]. These form the most rudimentary ethical design considerations for the IDCTA.

There are many frameworks through which the ethics of DCTA's use can be considered in greater detail. Upshur [48] described four guiding principles when considering whether a public health intervention is ethical (the harm principle, the principle of least restrictive means, the reciprocity principle, and the transparency principle). Childress et al [49] provide a set of five justificatory conditions (effectiveness, proportionality, necessity, least infringement, and public justifications) necessary for a public health intervention to interfere with individual liberties. Kass [50] described an ethical framework for public health interventions based on an assessment of the effectiveness, potential harms, alternative interventions and options for harm interventions of the intervention, and whether it is an equitable intervention. However, the complexity and lack of clarity of the ethical issues surrounding the relatively recent advent of digital contact tracing in the context of what was a new rapidly evolving global pandemic have necessitated frameworks that specifically address ethical considerations in digital contact tracing [15,51-56]. An overview of the key ethical considerations presented in these frameworks is provided in Multimedia Appendix 6 [15,51-56]. There is considerable agreement between the frameworks and among early guidance from the European Commission [57] and digital technology expert groups [14,58] about the considerations necessary to design and deploy DCTAs ethically, which broadly can be described as proportionality, voluntariness, transparency and trustworthiness, and equity.

Proportionality, that is, ensuring the intervention is a proportionate response to the public health threat, defines the ethical limits of other aspects of the IDCTA, such as its clinical and societal use and its interference with privacy and data protection rights [14,58,59,176]. There is no doubt that COVID-19 is a significant threat to public health, as evident even from early reports [177]. However, to determine if DCTAs are a proportionate response to this threat, an assessment of their potential benefits and risks is required. As a key component of the WHO-advised strategy to counter COVID-19 [5], any intervention that improves the effectiveness of contact tracing is of benefit to public health during the pandemic. However, high-quality evidence that DCTAs are effective in doing this is lacking [60,61]. Randomized controlled trials are the gold standard when evaluating the effectiveness of interventions, but for DCTAs, they may be logistically challenging and costly to design [62]. Augmenting manual contact tracing through nondigital means has proved problematic, particularly in Western societies [63,178]. In the absence of other alternatives to augment manual contact tracing and given the uncertainty regarding DCTA effectiveness, a key feature of the IDCTA should be harm minimization [53]. As summarized in Textbox 1, this includes minimizing the risk to personal data and privacy, and minimizing the risk of the false characterization of contact status.

Textbox 1. Potential risks of digital contact tracing apps.

Data protection

- Failure to protect personal data from misuse [64,65]

Privacy

- Loss of personal privacy with no personal or societal benefit [59,61,66,176]

Resources

- Misuse of limited financial and human resources on an ineffective intervention [67,68]

Clinical

- False-positive characterization of contact status (may result in unnecessary quarantining and anxiety) [69,70]
- False-negative characterization of contact status (may result in further onward disease transmission) [69,70]

Public engagement

- Loss of trust in public health authorities and public health measures [14,56,66]

The IDCTA should be voluntary and consent-based according to the WHO and ECDC ethical guidance on DCTAs, and this view is also prevalent in the academic literature [40,51-55]. For voluntary DCTAs, ensuring transparency and trustworthiness is important to maximize population penetration [52,53,71-79]. Transparency can be achieved by making the underlying DCTA algorithms open source [52,80,81], and this is established as best practice [82,83]. Similarly, an explanation of the risk prediction algorithm used should be publicly available for scrutiny [52] as has been done by Germany's Corona-Warn-App [179]. The data processed by the IDCTA, including the type, purpose, and duration of data storage, should be presented in an accessible, transparent way for all DCTA users [52,80,81]. Trustworthiness (how trusted a DCTA is by the target population) depends not only on transparency, robust data protection, and privacy preservation [37,72,84,85] but also on having published guiding ethical principles [86], a defined timeline and plan for DCTA evaluation, a defined published criteria for DCTA deactivation, and a DCTA independent oversight committee that has representation from civic society and the public [15,51,52].

DCTA availability and accessibility must be equitable, and they should not be used in a discriminatory way [52,56,66,87,175]. This is important not only from a human rights perspective but also from an effectiveness perspective because the success of a DCTA depends on factors such as user penetration within the general population [88] and in high-risk population groups [15,89]. The IDCTA should be disseminated free of charge, so it is accessible by all societal groups but, in particular, those disproportionately affected by COVID-19 such as older adults, people of lower socioeconomic class, and ethnic minorities [51,52,56,61,90,91]. Smartphone technology may be inaccessible to people of low incomes or those with limited digital literacy [92]. In Singapore, Bluetooth-enabled contact tracing tokens have been distributed to older adults who were less likely to be smartphone owners [93]. This practice should be encouraged, and alternatives such as free smartphones or monofunctional digital contact tracing devices should be deployed in parallel to groups who may not otherwise have access to DCTAs [14,94,95].

User Experience Considerations

The IDCTA should be designed so that it synchronizes two independent environments, that of public health authorities and that of end users. DCTA user experience considerations can be thought of as those relating to universality and those relating to user engagement. Multimedia Appendix 7 [29,32,36,38,39,42,69,96-104,106,107,110-118,180-182] provides an overview of the key academic literature [75,96-105], gray literature [106,107], regulations [42,108-111], guidelines [112,113,180], and assessments of existing DCTAs [29,32,38,114-118] that support these recommendations.

To support a more holistic approach to the design of the IDCTA, the concept of universality allowed the cross-disciplinary team to identify a series of dimensions to be taken into account, such as accessibility, minors as users, cultural universality, content, availability, and maintenance and frequency of upgrades with the aim of better accommodating different users' needs, including minors, older adults, people with chronic disease, and those with various forms of disability, so that accessibility and inclusiveness can be ensured [42]. In keeping with this holistic approach to the design of the IDCTA, the interface elements should enable multimodal interaction (eg, supported by voice control) with contents that are available in different languages. Additionally, jargon should be avoided. How well the dimensions of universality are incorporated into the design of the IDCTA will affect not only its population penetration and continued use but also its interoperability across borders. Population penetration will depend on the prevalence of smartphones and operating systems in use among the target population that support the chosen DCTA technology [75,105]. Trade-offs between accuracy and availability will need to be assessed so that it is available on the widest range of smartphones and operating systems possible (eg, ultra-wideband is accurate but not widely available [183,184] and Bluetooth LE is less accurate but more widely available [119,120,185]). This must be done while also supporting the various screen

sizes and resolutions of the widest range of smartphones in use among the population. The IDCTA should be conceived as open-ended with frequent updates, ongoing support, and constant maintenance.

Regarding user engagement, nine key aspects were identified from across the literature that could help improve engagement: performance feedback, helpfulness, public health measures, educational information, personal information, personalization and control, time and human effort, flexibility or multimodality, and multitasking. Based on these aspects, user requirements that could increase engagement are evident. Engagement could be potentially enhanced by enabling the user to contact their case health care worker should they have questions regarding their COVID-19 diagnosis. Engagement could also potentially be enhanced by allowing users to identify areas where the incidence of COVID-19 infection is high that they may wish to avoid or settings where the risk of contracting COVID-19 when exposed may be highest (eg, public transport routes known to be frequently crowded). Dynamic, consistently updated information on confirmed cases, testing sites, vaccination sites, government restrictions, and preventive strategies could enhance user engagement by making the benefit of using the app more apparent to the users and integrating it with the wider public health information campaign as part of the national COVID-19 response. However, the amount of information presented should not be overwhelming for users. Graphic representation of these data may also be beneficial (eg, visualization may summarize the number of cases or close contacts being reported per day or week). By conveniently providing useful information on the DCTA, it has the potential to engage and help users long-term to protect themselves against COVID-19. The IDCTA should also enable the end user to tailor the app to their particular needs to enhance user engagement. For example, users might find it beneficial to personalize which notifications they receive or to temporarily deactivate the contact tracing function [121].

Privacy and Data Protection Considerations

Privacy and data considerations of DCTAs are dependent on what their functional requirements are. DCTAs need to maintain a contact log, generate a contact alert, and link users with the test and trace system. The IDCTA needs to perform these functions while respecting individual privacy rights and adhering to data protection regulation [14,22,37,40-42,52,56,69,72,74-76,84,122-124]. The European Charter of Human Rights Article 8 states that individuals have a right to respect for private life, but interference with this right can occur if it is deemed necessary, proportionate, and in accordance with the law [186]. The IDCTA should follow the foundational principles of privacy by design, a widely used approach in systems engineering characterized by proactive rather than reactive measures, and this approach to digital contact tracing is supported by the European Data Protection Board (EDPB) [125]. The collection and use of personal data are protected by several regulations in the European Union, such as the European General Data Protection Regulation (GDPR) Act 2016 and the ePrivacy Directive 2002 [126,187]. Article 6 of the GDPR states that processing of data is lawful if it is “necessary to protect the vital interests of a person” and if it is “necessary for the performance of a task carried out in the public interest” [187]. Contact tracing of infectious diseases is lawful because it is necessary to protect case contacts, and epidemic containment is certainly carried out in the public interest, a view which is supported by the WHO and ECDC [5,40]. Data concerning health such as one’s COVID-19 infection or contact status are considered “special data” as described in Article 9 of the GDPR [187]. The processing of “special data” is permissible only where “processing is necessary for reasons of substantial public interest,” which shall be “proportionate to the aim pursued” and respects “the essence of the right to data protection” [187]. More specifically, special data may be processed if it is “necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care” [187]. As further explained in recital (46) of the GDPR, the processing of personal data should also be regarded to be lawful where it is necessary to protect an interest that is essential for the life of the data subject or that of another natural person, for instance, when processing is for monitoring epidemics [127]. Evidently, data collection and processing to facilitate contact tracing during an epidemic to prevent further disease transmission and death is permissible. There are limits to this as defined in Article 5 of the GDPR, which sets out seven key principles related to the processing of personal data (Textbox 2) [128].

Textbox 2. Principles of data protection.

Lawfulness, fairness, and transparency

- Lawfulness: Processing of personal data carried out by a controller must have a legal basis under the General Data Protection Regulation.
- Fairness: Processing of personal data must be fair toward the individual whose personal data are concerned and avoid being unduly detrimental, unexpected, misleading, or deceptive.
- Transparency: Controllers must provide individuals with information regarding the processing of their personal data in a format that is concise, easily accessible, and easy to understand.

Purpose limitation

- Personal data must be collected for specified, explicit, and legitimate purposes.

Data minimization

- Personal data that are collected and processed should be adequate, relevant, and limited to what is necessary for the purposes for which they are processed.

Accuracy

- Personal data that are collected should be accurate and, where necessary, kept up to date.

Storage limitation

- Controllers must hold personal data, in a form that permits the identification of individuals, for no longer than is necessary for the purposes for which the personal data are processed.

Integrity and confidentiality

- Personal data must be processed by controllers only in a manner that ensures the appropriate level of security and confidentiality for the personal data using appropriate technical or organizational measures.

Accountability

- Controllers are responsible for, and must be able to demonstrate compliance with, the other principles of data protection.

To be lawful, there must be a legal basis on which the data are processed [109]. For example, in the ROI, a DCTA was introduced on the legal basis set out under section 7 of the Health Act 2004, which states that health authorities should use its resources to protect the health and welfare of the public [185]. Use of the IDCTA should be voluntary, and this should be included in legal frameworks when legislating for its use [40,51-55]. To ensure accountability, the controller of the DCTA should be clearly defined and the EDPB suggest this could be national public health authorities [41].

DCTA contact logs should adhere to the principles of privacy by design and data minimization by collecting only an anonymized identifier unique to each contact event [41]. This means the IDCTA should not record the name, age, sex, ethnicity, or address of the contact nor should it record the time or location of the contact event [41]. However, privacy needs to be embedded into the DCTA design without diminishing functionality as much as is possible [125]. To enable risk stratification of the contact event, the IDCTA should record the day of the contact event, as is done by the Corona-Warn-App [179]. Although collection of location data is recommended against by the EDPB, in the ROI and the United Kingdom, there is some evidence that the majority of people do not have an objection to its use by DCTAs in the context of an epidemic [37,121,129]. Although this may vary between countries, where location tracking is deemed a proportionate response to the scale of the epidemic, the IDCTA should log the location of the app user locally [121] but not that of their contacts in keeping with privacy by design principles [125]. The principle of purpose limitation [128] would dictate that both contact and location tracking logs should be collected only for the purpose of COVID-19 contact tracing. How a contact event is recorded by a DCTA must be accurate; otherwise, there is the potential for large-scale misclassification of contact events as occurred with the UK NHS COVID-19 app [34,35]. Therefore, field studies that validate the accuracy of the app should be performed and published as has been done for some DCTAs [130]. In the ROI, independent assessments of COVID Tracker have been performed and published, and this practice should be encouraged [131,132]. Contact logs should be maintained for 14 days, the incubation period of COVID-19 [3], to adhere to the principle of storage limitation.

When a DCTA user is confirmed to have COVID-19, an exposure notification system is necessary to enable them to alert their contacts. To ensure data collected are accurate, it has been suggested that COIVD-19 cases have their status verified before they can use the exposure notification system to prevent misuse [133,134]. Verification should preferably be automated [135]. To ensure integrity and confidentiality, the EDPB [41] and European Union eHealth Network [42] recommend that contact log processing follow a decentralized privacy preserving protocol (ie, processing of contact logs to match those of the user's contacts with those of cases that occurs on the user's device). The principle of privacy by design [125] would dictate

that contact alerts should be generated from the contact's DCTA as opposed to being sent from the case's DCTA. The contact alert, to be adherent to the principle of data minimization, should not contain the cases' personal information such as name, age, sex, or ethnicity nor should it contain the time or location of the contact event [41]. For example, contact alerts generated by COVID Tracker in the ROI inform users "Close Contact Alert: The app has detected that you have been in close contact with someone who has tested positive for COVID-19" [29]. When a contact alert is generated, the DCTA user should be able to contact public health authorities through the app [32], or they should be provided with a number to contact health authorities on.

In some countries such as South Korea and Israel, interference with individual privacy rights was deemed to be a necessary proportionate response to COVID-19. However, Western societies particularly value privacy [37,136]. Maximizing population penetration of a voluntary DCTA in these societies will require health authorities to convince target users that their privacy will be protected [68,72,77,137]. The nature of data collected, whether it be proximity data, location data, or both, hinges on whether it is deemed to be proportionate to the aim pursued. The EDPB state that DCTAs should rely on proximity and not location data [41]. Many countries such as the ROI, the United Kingdom, and Germany have developed DCTAs that record proximity using Bluetooth LE [29,34,179]. However, from the experiences of Israel and South Korea, location tracking may be a key feature of effective digital contact tracing. Tracking location may be useful to identify previously unknown settings where transmission is occurring, allowing for public health authorities to take proactive action to prevent further transmission [25,138,139]. However, the privacy risks are significant, and misuse of location data can be harmful to public trust in health authorities [23]. There is a need for public engagement mechanisms in each country to define by consensus what the limits of a proportionate response to COVID-19 are. Where location tracking is used, it should be an opt-in feature [121] because invasion of personal privacy can be a significant deterrent to downloading and using a DCTA, and the use of geolocation data has recently been the cause of privacy losing events associated with app use [85,140].

In May 2020, Google and Apple collaborated to create an application programming interface (API) [141]. An API is a *Lego block* on which governments can build a DCTA. The DCTA (interface, data collection, public health information), server (epidemiological dashboard, diagnosis verification), and server relay (for 14 days of the case's cryptogenic keys) are supplied by the public health authority [141]. The Google/Apple API works on Android 6 and iOS 13 forward [141]. Building a DCTA on this API requires a series of measures to protect individual privacy. There should be a requirement for explicit user consent, anonymity of all users to each other, and allowing users the choice of how much personal information they share [81,141]. Google and Apple control which DCTAs use the API and for how long the API is operational in a given region [141]. There should be an agreed timeline and criteria for when the API and DCTA infrastructure is to be dismantled [76,121,124,141] because there are significant privacy concerns that technology companies and governments could use DCTAs to enable greater surveillance after the pandemic [26,37,65,66,76]. Using an API comes with the risk that personal data may be misused or processed unlawfully. Google has a record of not adhering to data protection regulations and the principles of data protection [188]. There are also ongoing concerns regarding Google's lawful and transparent use of location data [189]. The entry of private corporations into pandemic response may create a dependency on them to deliver public health necessities, global health policies, and result in an accumulation of decision-making powers across multiple aspects of society and subversion of democratically elected governments [59,176]. Despite concerns surrounding the role of private corporations in digital contact tracing, the number of app downloads was high [37,190]. This may be explained by the phenomenon known as the privacy paradox, whereby people express concerns regarding sharing personal information, but their behavior is incongruent with the concerns they express [142].

Technical Considerations

For the IDCTA, the choice of which technology to use to detect contact events will be influenced by its cost, energy use, accuracy, availability, accessibility, adherence to data protection regulation, and by how it effects privacy preservation and overall DCTA effectiveness [14,17,37,40,42,52,61,69,76,141]. These in turn influence DCTA population penetration. Potential technologies include ultra-wideband technology, Wi-Fi, Bluetooth LE, ultrasound, and GPS (Table 1). Ultra-wideband technology is an ideal technology for proximity detection. It is a low cost, low energy use technology that can measure highly accurate spatial data both indoors and outdoors, with an ability to discriminate distances of 10 to 30 cm [143,183,184]. With regard to privacy, it is considered more secure than Bluetooth LE [143]. However, smartphones equipped with this technology are not yet in common use [184]. A DCTA that uses Wi-Fi would be limited by range and difficult to make ubiquitously available. According to the EDPB and the European Union eHealth guidance, Bluetooth LE proximity detection should be used because it maximizes privacy preservation and is widely available, which are important to maximize population penetration [41,42,119]. However, population penetration will also rely on belief in the accuracy of the DCTA to detect contact events [79,105,144,145]. DCTAs that use Bluetooth LE alone may not have adequate accuracy [70,132,146,147,185]. GPS location tracking may be less accurate indoors or in multistory buildings as compared with Bluetooth LE [44,148]. Published studies validating the accuracy of these DCTA technologies are lacking (Table 1). Combining ultrasound technology with Bluetooth LE may improve accuracy by reducing the number of false-positive contacts identified [116,130]. In the absence of widely available ultra-wideband technology, this represents the best compromise on privacy preservation, accuracy, and availability for the IDCTA. However, this may not be applicable across all countries because the choice of technology will depend on the prevalence of compatible smartphones in use in the population and how valued personal privacy is among the population.

Table 1. Potential DCTA technologies and the implications of their use.

DCTA[a] technology	Bluetooth LE[b]	GPS-enabled geolocation tracking	Bluetooth LE and ultrasound	Ultra-wideband
Accuracy[c]	Accuracy reported as 72%[d] (distance threshold not reported) and 79% (distance threshold 1.5 m); although, independent studies did not reproduce these results [70,147,185].	Accurate to within 4.9 m, but concerns that GPS location tracking for COVID-19 contact tracing not feasible due to limited accuracy [149,191]	Accuracy reported as 55% (distance threshold ≤6 foot) and accuracy reported as 99.6% (distance threshold ≤12 foot) [130]	Highly accurate [143]
Effectiveness in augmenting manual contact tracing	Limited evidence to suggest effectiveness [150,151]	Limited anecdotal evidence to suggest effectiveness [25,138]	Insufficient evidence found to suggest effectiveness	No instances of ultra-wideband–enabled DCTAs found in the literature.
Energy use	Less than GPS [148]	More than Bluetooth LE [148]	Not reported	Low energy use [143]
Accessibility and availability	Widely available [119,120]	Widely available [120]	Widely available but less so than Bluetooth LE or GPS	Not widely available [143,183,184]
Adherence with principle of privacy preservation	Highly adherent (records only proximity)	Less adherent (records location, which is potentially identifiable)	Adherent[e] (records only proximity)	Highly adherent [143] (records only proximity).
Adherence with principles of data protection	Adherent	Interferes with the principle of data minimization	Adherent	Adherent

[a]DCTA: digital contact tracing app.

[b]LE: Low Energy.

[c](True positives + true negatives) / total number of tests.

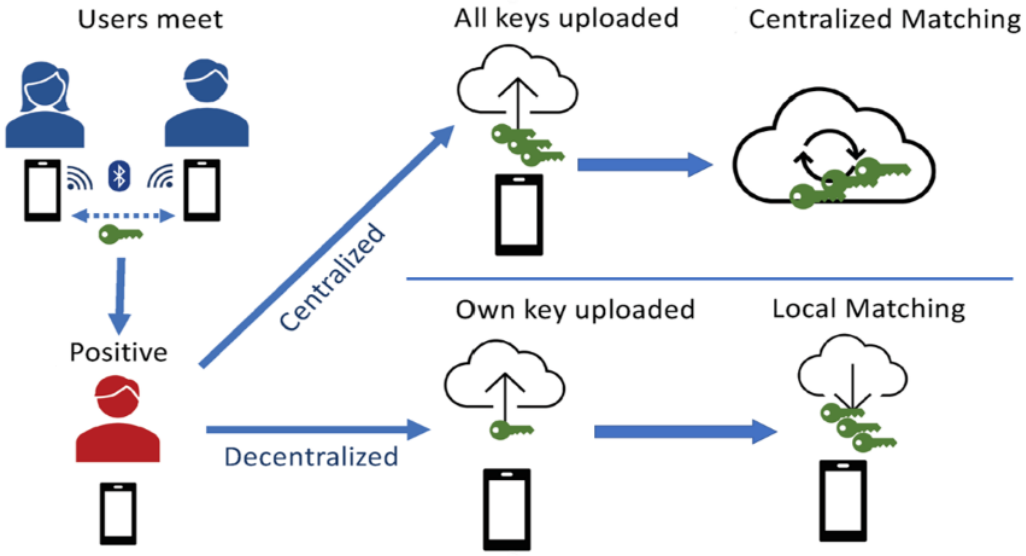
[d]COVID Tracker Ireland reported being able to accurately identify 72% of close contacts, although field studies supporting this claim have not been published.

[e]Perception that it has the potential for misuse of audio data [152], but this is not the case according to proponents of this technology [153].

The IDCTA must enable users to exchange and record temporary contact numbers when they are in contact within prespecified time and distance thresholds [44,154]. Temporary contact numbers should be renewed frequently (eg, every 15 minutes) to protect user privacy [44,141,154]. Contact logs of temporary contact numbers are maintained by each device, and once a case of COVID-19 is diagnosed, the DCTA allows them to notify their contacts [44,141]. How contact logs are processed to identify contacts and how contacts are alerted can be performed in a centralized or decentralized manner (Figure 2). Early in the pandemic, two predominant protocols emerged, the Pan-European Privacy Preserving Proximity Tracing (PEPP-PT initiative; centralized) and the Decentralized Privacy Preserving Proximity Tracing (DP-3T initiative; decentralized) [155,156]. With the PEPP-PT initiative, contact logs from the case's device are processed centrally, while with the DP-3T initiative, contact logs are processed on the contact's device by regularly checking a central server that holds the temporary contact numbers of cases [155,156]. The DP-3T initiative provides more protection to individual privacy and may enhance DCTA uptake [155,157]. The PEPP-PT initiative involves a human-in-the-loop, which is disadvantageous because it shares the case's contact log with another individual. However, this may minimize false positives related to contact occurring through apartment walls or where adequate contact precautions were in place [156]. Centralized collection of personal data leaves individuals vulnerable to social network mapping and potentially having their movements mapped. Decentralized protocols may also be vulnerable to malicious attacks [133]. A user's temporary contact number could be accessed and used by multiple devices and result in false contact chains being generated during the life cycle of that temporary contact number.

Ensuring processes that protect personal data from misuse are rigorously enforced is important in building public confidence that their personal data are safe. The European Union has stated that a DCTA should use a decentralized model to protect individual privacy [42]. They also emphasized that DCTAs should augment and not replace existing contact tracing systems [42]. Automated contact alert notification using a decentralized protocol may be alarming for contacts to receive. Therefore, the IDCTA should not only follow a decentralized privacy preserving protocol but also provide explicit instruction on what actions to take if a contact alert is received and a means of making contact with a health care worker for integration into the test and trace system.

Figure 2. Centralized versus decentralized digital contact tracing. Reprinted from Hernández-Orallo et al [192] under the Creative Commons CC-BY 4.0 license.



How a contact is defined by the app should also be considered. DCTAs may define contacts according to binary distance (eg, within 2 m) and duration of contact (eg, 15 minutes or less) thresholds in keeping with the definition applied by health authorities [8,158,159]; although, this binary definition may not identify all contacts at risk of infection [160]. Alternatively, a risk-stratified approach that includes other factors associated with disease transmission may identify contacts more accurately [34]. Contact could be stratified as *high risk* or *low risk* depending on a risk score, with high-risk contacts being advised to quarantine and arrange COVID-19 testing and low-risk contacts being advised on good social distancing practice. Risk scores could be based on how close the contact was, for how long the contact lasted, how long it has been since the user met a COVID-19–positive person, and the risk of transmission for the case [161]. The risk of transmission would be based on empirical evidence of COVID-19 transmission dynamics. Temporary contact numbers recorded in the case's DCTA each day would be assigned an additional code that represents the risk value of COVID-19 transmission from the case on that day. The contact's DCTA would, in addition to recording the temporary contact number, record this additional code so that it may calculate the COVID-19 risk score when it matches temporary contact numbers from its contact log with that of a case. Risk scores are only as accurate as the data they are constructed from, and defining the risk score threshold would require an ongoing process of evaluation and calibration [42]. To calibrate the DCTA risk score effectively, knowledge of the contact event outcomes would be needed. To do this, the IDCTA would allow users to voluntarily have their COVID-19 test result uploaded directly to their DCTA from the processing laboratory (as is possible with the Corona-Warn app [115]) or for users to voluntarily confirm their status as having COVID-19 on the DCTA (as occurs with COVID Tracker [29] and COCOA [134]). The DCTA would have to centrally collect for users who receive a contact alert and who volunteer their COVID-19 test result, how long ago the contact occurred, the duration of contact, and the proximity of contact.

Clinical and Societal Considerations

COVID-19 transmission chain disruption could potentially be enhanced by using DCTAs to augment manual contact tracing [16]. Therefore, this is the aim of the IDCTA. Other aims may be to act as a confidence-enhancing measure for those most vulnerable to COVID-19 infection. It may reassure them that they have not been a contact, and others can demonstrate their contact-free status to them [162]. The higher the uptake of the IDCTA, the more likely it is a case can notify a contact of their exposure. To achieve this, demonstrating to the target population the high degree of privacy preservation and adherence to data protection regulation is important. Enhancing uptake through small monetary incentives may also be considered [163,164]. The IDCTA should be voluntary [40,51-55,134], and mandating the use of DCTAs of uncertain effectiveness with associated potential harms, which have the highest likelihood of utility during periods of low COVID-19 incidence, may be difficult to justify as a proportionate response.

The IDCTA should avoid functions that necessitate additional data processing that may raise privacy concerns, such as age, sex, location, or ethnicity. Any additional functions should be justifiable, proportionate, privacy preserving, and adherent to data protection regulation. Additional functions should be defined before DCTA deployment in keeping with the data protection principle of purpose limitation [128]. DCTAs present an opportunity to perform functions such as allowing people to assess their personal risk of being hospitalized or dying from COVID-19. Providing risk assessments may not be ethical given that risk algorithms may be population specific, not generalizable, and may provide falsely elevated or falsely lowered risks [165]. Regardless of risk, it could be argued that the person's behavior should be the same and an awareness of one's risk may not result in positive behavior change. Therefore, the IDCTA would not provide risk assessments. Symptom checker functions should also be avoided because there is little high-quality evidence to support their use in this context.

Evaluation Considerations

To be ethical and adherent with data protection regulation, the continued use of a DCTA needs to be supported by evidence that it has been effective in contributing to epidemic control. A DCTA is a multistep intervention. There are several steps where they may fail to effectively disrupt transmission chains, including being downloaded; recording contact events; sending contact alerts; and integrating with the wider contact tracing, testing, quarantine, and isolation systems [16,166]. Maintaining privacy while ensuring the necessary data to demonstrate effectiveness are collected is challenging. To enable evaluation of effectiveness, the IDCTA should record and collect key metrics (Table 2). These were derived from ECDC guidance on how to monitor contact tracing effectiveness [159], the limited number of studies evaluating real-world DCTA effectiveness [150,151,167-170], and other published academic literature [62]. Many of these metrics may be collected by the app and do not interfere with individual privacy. Other metrics such as the outcome of COVID-19 testing would be considered sensitive data by many. Collection of these sensitive data should be voluntary. Determining whether people who receive contact alerts quarantine, a key intervention in disrupting transmission changes, may be difficult. This is true for both digital and manual contact tracing. GPS location tracking has been used in China by public health authorities to confirm contacts remain within quarantine [171]. However, this interference with the right to privacy would likely not be acceptable in countries where privacy is highly valued. The outcomes of the evaluation should be openly reported to allow for a wider public evaluation of the app.

Table 2. Metrics to evaluate ideal DCTA effectiveness.

Indicator of effectiveness	Purpose	Metric numerator (source)	Metric denominator (source)
DCTA[a] is downloaded	To estimate the proportion of the smartphone owning population who download the DCTA	Number of DCTA downloads minus number of DCTA deletions (DCTA)	Number of smartphone owners nationally (Government statistics office; eg, Central Statistics Office, ROI[b])
DCTA is active	To estimate the proportion of DCTAs downloaded that are being used	Number of DCTAs with contact tracing turned on (DCTA)	Number of DCTAs downloaded minus number of DCTAs deleted (DCTA)
DCTA is active	To estimate the proportion of DCTAs downloaded that are being used	Frequency and duration of use (DCTA)	N/A[c]
DCTA is active	To estimate the proportion of DCTAs downloaded that are being used	Number of DCTAs downloading TCNs[d] of cases on central server per day (assuming DCTA downloads keys once per day when active; DCTA)	Number of DCTAs downloaded minus number of DCTAs deleted (DCTA)
DCTA is used by COVID-19 cases	To estimate the DCTA penetration among people who contract COVID-19	Number of positive test results uploaded to DCTA (DCTA)	Number of COVID-19 cases nationally (national surveillance data)
DCTA is used by COVID-19 cases	To estimate the DCTA penetration among people who contract COVID-19	Number of COVID-19 cases who attended a screening center reporting DCTA active use (survey of attendees at testing centers and review of participants' test results)	Number of COVID-19 cases who attended a screening center (screening center data)
DCTA is used by COVID-19 cases to notify close contacts	To estimate the proportion of cases using the DCTA who use it to send contact alerts	Number of DCTAs that send a contact alert (DCTA)	Number of DCTAs with a positive COVID-19 test recorded (national surveillance data)
DCTA is used by COVID-19 cases to notify close contacts	To estimate the proportion of cases using the DCTA who use it to send contact alerts	Number of COVID-19 cases who attended a screening center reporting DCTA active use and who report sending a contact alert (follow-up survey of COVID-19 cases who reported DCTA use at time of screening)	Number of COVID-19 cases who attended a screening center reporting DCTA active use (survey of attendees at testing centers and review of participants' test results)
Close contacts using DCTA receive alert	To estimate the DCTA penetration among people who are close contacts	Number of DCTAs that receive a contact alert (DCTA)	Number of close contacts identified nationally (national surveillance data)
DCTA identifies contacts not identified by manual contact tracing	To demonstrate the DCTA augments manual contact tracing	Number of close contacts attending testing center identified exclusively by DCTA (survey of attendees at testing centers)	Number of close contacts attending testing center (survey of attendees at testing centers)
DCTA identifies contacts sooner than manual contact tracing	To demonstrate the DCTA augments manual contact tracing	Number of close contacts attending testing center who received contact alert from DCTA before contact alert from manual contact tracing service (survey of attendees at testing centers)	Number of close contacts attending testing center (survey of attendees at testing centers)
Close contacts using DCTA are tested for COVID-19	To estimate the proportion of contacts who are tested for COVID-19 and to estimate the number of cases identified by the DCTA	Number of DCTAs with a COVID-19 test result uploaded within 14 days of a contact alert (DCTA)	Number of DCTAs that receive a contact alert (DCTA)
DCTA associated harm is recognized	To determine what harms, if any, occur with DCTA use	N/A (qualitative survey of DCTA users)	N/A

[a]DCTA: digital contact tracing app.

[b]ROI: Republic of Ireland.

[c]N/A: not applicable.

[d]TCN: temporary contact number.

Summary of Findings

Key considerations were ethical, user experience, privacy and data protection, clinical and societal, and evaluation. Proportionality, voluntariness, transparency, trustworthiness, and equity are necessary for the design and deployment of the IDCTA. Universality and user engagement are important user experience considerations that can influence DCTA use in the

population. Dimensions of universality that should be taken into account when designing the IDCTA are accessibility, minors as users, cultural universality, content, availability, and maintenance and frequency of upgrades. User engagement could be enhanced by small financial incentives, enabling users to tailor aspects of the app to their particular needs and integrating DCTAs into the wider public health information campaign. If DCTAs are to be trusted, accepted, and used by the target population, they must be adherent to data protection regulation and have privacy by design through all elements, including maintaining contact logs, generating contact alerts, and linking users into the test and trace system. For the IDCTA, the choice of which technology is used will be influenced by its cost, energy use, availability, accessibility, adherence to data protection regulation and principles of privacy by design, and accuracy when detecting contact events. Combining ultrasound technology with Bluetooth LE may improve accuracy by reducing the number of false-positive contacts identified. A decentralized privacy preserving protocol should be followed to enable DCTA users to exchange and record temporary contact numbers during contact events. The IDCTA should define and risk stratify contact events according to proximity, duration of contact, and the infectiousness of the case at the time of contact. Evaluating DCTAs requires data to quantify app downloads, use among COVID-19 cases, successful contact alert generation, contact alert receivers, contact alert receivers that adhere to quarantine and testing recommendations, and the number of contact alert receivers who subsequently are tested positive for COVID-19.

Discussion

Principal Findings

This cross-disciplinary review presents best practice guidance for developing the IDCTA and is informative for those involved in DCTA research, design, and deployment. It also serves as a comprehensive and accessible entry point for those beginning to engage with this research subject, which has evolved significantly after a period of intensive exploration in 2020. DCTAs will likely be a significant research field not only for the remainder of the COVID-19 pandemic but also in the postpandemic era because of a renewed interest and support for pandemic preparedness. Demonstrating the effectiveness of COVID-19 DCTAs is a current research priority [193]. This is important to convince not only nonapp users of their benefits but also current or previous app users, many of whom remain uncertain about their utility [194]. Early evidence indicates that DCTAs can identify contacts of COVID-19 cases who subsequently develop infection (particularly among nonhousehold contacts) [150,151,167-170], may shorten the time to quarantine by 1 day [168], and can prevent further disease transmission [169]. Ensuring DCTAs are integrated with the wider test and trace system is emerging as an important aspect of DCTA deployment [135]. Where codes were required by DCTA users to confirm on the app a COVID-19–infected status, manual distribution of these codes by health care professionals could delay contact alert generation and subsequent downstream actions such as contact quarantine and testing, suggesting automated code generation is preferable [135].

A weakness of this research was that it did not specifically address how DCTAs should be integrated into the wider test and trace system. There is a need for future dedicated research to synthesize and evaluate evidence, and generate best practice recommendations for this consideration of DCTA deployment. The limitations of this review are that the index and gray literature searches, while extensive, were not performed using systematic review methodology. The inclusion of both indexed and gray literature enabled the derivation of best practice guidance from the literature during a phase of rapid DCTA research and development growth. The cross-disciplinary approach taken to evaluating the evidence was a strength of this research because it allowed varying aspects of DCTA design and deployment to be considered.

Future promising developments in this field may be the use of blockchain technology, ultra-wideband technology, and artificial intelligence in DCTA design. Privacy and data protection concerns are significant barriers to DCTA uptake in Western societies [72,74-76,122-124]. A blockchain network is a decentralized, distributed, and secure public ledger that stores records of transactions securely using cryptography techniques [195]. Features of blockchain technology that make it advantageous for digital contact tracing are decentralized data storage; data security through encryption; data provenance and time stamping allowing for verification of the data legitimacy and data immutability, which enhances data reliability and transparency [196]. The use of blockchain networks in future DCTAs may reduce privacy and data protection concerns and enhance DCTA uptake and use [196]. This area should be a focus of future research. DCTAs need to detect contact events accurately to optimize uptake [79,105,144,145]. Ultra-wideband is a low energy means of enabling short-range high bandwidth communications that can transmit data with minimal noise interference. This could allow for highly accurate measurement of contact events within centimeters [197]. Although not yet a feature of most smartphones, it most likely will be in the near future [198]. Therefore, it could be a viably accessible and available energy efficient technology for DCTAs in future pandemics. Additionally, artificial intelligence could potentially improve the accuracy of future DCTA contact event detection by reducing false positives and false negatives [199].

Conclusion

In conclusion, key considerations and best practice guidance for the design of the IDCTA were derived from the literature.

Acknowledgments

This study was supported by Science Foundation Ireland grants 20/COV/0133 and 13/RC/2094.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cross-disciplinary approach.
[DOCX File , 228 KB - mhealth_v9i6e27753_app1.docx]

Multimedia Appendix 2
Scoping review.
[DOCX File , 15 KB - mhealth_v9i6e27753_app2.docx]

Multimedia Appendix 3
Search strategy.
[DOCX File , 13 KB - mhealth_v9i6e27753_app3.docx]

Multimedia Appendix 4
Gray literature sources.
[DOCX File , 21 KB - mhealth_v9i6e27753_app4.docx]

Multimedia Appendix 5
Included articles' description.
[DOCX File , 79 KB - mhealth_v9i6e27753_app5.docx]

Multimedia Appendix 6
Ethical frameworks for digital contact tracing apps.
[DOCX File , 18 KB - mhealth_v9i6e27753_app6.docx]

Multimedia Appendix 7
Evidence supporting user experience recommendations.
[DOCX File , 30 KB - mhealth_v9i6e27753_app7.docx]

References

1. Timeline: WHO's COVID-19 response. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline [accessed 2020-09-26]
2. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2021-05-28]
3. Coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses [accessed 2020-09-26]
4. Coronavirus disease 2019 (COVID-19): situation report, 46. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/331443 [accessed 2021-03-17]
5. Critical preparedness, readiness and response actions for COVID-19: interim guidance. World Health Organization. 2020 Jun 24. URL: https://apps.who.int/iris/handle/10665/332665 [accessed 2020-09-26]
6. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 May;26(5):672-675. [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
7. Barnett-Howell Z, Mobarak AM. The benefits and costs of social distancing in rich and poor countries. arXiv. Preprint posted online on April 10, 2020.
8. Contact tracing in the context of COVID-19: interim guidance. World Health Organization. 2020 May 10. URL: https://apps.who.int/iris/handle/10665/332049 [accessed 2020-09-26]
9. Resource estimation for contact tracing, quarantine and monitoring activities for COVID-19 cases in the EU/EEA. European Centre for Disease Prevention and Control. 2020. URL: https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-resources-for-contact-tracing-2-March-2020_0.pdf [accessed 2020-09-26]
10. Public Health England. Guidance for contacts of people with confirmed coronavirus (COVID-19) infection who do not live with the person. GOV.UK. URL: https://www.gov.uk/government/publications/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person [accessed 2020-11-03]
11. Novel Coronavirus 2019 (COVID-19) National Interim Guidelines for Public Health management of contacts of cases of COVID-19. Health Protection Surveillance Centre. 2020. URL: https://www.hpsc.ie/a-z/respiratory/coronavirus/

novelcoronavirus/guidance/contacttracingguidance/National%20Interim%20Guidance%20for%20contact%20tracing.pdf [accessed 2020-09-26]
12. Linton N, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov A, Jung S, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. J Clin Med 2020 Feb 17;9(2):538 [FREE Full text] [doi: 10.3390/jcm9020538] [Medline: 32079150]
13. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis 2020 Apr;93:284-286 [FREE Full text] [doi: 10.1016/j.ijid.2020.02.060] [Medline: 32145466]
14. Exit through the App Store? Ada Lovelace Institute. 2020 Apr. URL: https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-Exit-through-the-App-Store-Explainer-for-Government-April-2020.pdf [accessed 2021-02-04]
15. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
16. Kretzschmar M, Rozhnova G, Bootsma M, van Boven M, van de Wijgert J, Bonten M. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. Lancet Public Health 2020 Aug;5(8):e452-e459 [FREE Full text] [doi: 10.1016/S2468-2667(20)30157-2] [Medline: 32682487]
17. Bradshaw WJ, Alley EC, Huggins JH, Lloyd AL, Esvelt KM. Bidirectional contact tracing dramatically improves COVID-19 control. medRxiv. Preprint posted online on July 14, 2020. [doi: 10.1101/2020.05.06.20093369]
18. Hinch R, Probert W, Nurtay A, Kendall M, Wymant C, Hall M, et al. Effective configurations of a digital contact tracing app: a report to NHSX. GitHub. 2020 Apr. URL: https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Report%20-%20Effective%20Configurations%20of%20a%20Digital%20Contact%20Tracing%20App.pdf [accessed 2021-05-21]
19. Woodhams S. COVID-19 digital rights tracker. Top10VPN. URL: https://www.top10vpn.com/research/investigations/covid-19-digital-rights-tracker/ [accessed 2020-11-26]
20. What is the TraceTogether Programme? TraceTogether. URL: https://support.tracetogether.gov.sg/hc/en-sg/articles/360053530773-What-is-the-TraceTogether-Programme-How-is-it-different-from-the-TraceTogether-App- [accessed 2020-11-12]
21. 20 April 2020 - one month on. TraceTogether. URL: https://support.tracetogether.gov.sg/hc/en-sg/articles/360046475654-20-April-2020-One-Month-On [accessed 2020-11-12]
22. COVID-19 digital apps need due diligence. Nature. 2020. URL: https://media.nature.com/original/magazine-assets/d41586-020-01264-1/d41586-020-01264-1.pdf [accessed 2021-02-04]
23. Kim N. 'More scary than coronavirus': South Korea's health alerts expose private lives. The Guardian. URL: https://www.theguardian.com/world/2020/mar/06/more-scary-than-coronavirus-south-koreas-health-alerts-expose-private-lives [accessed 2020-11-12]
24. Holmes O. Israel to track mobile phones of suspected coronavirus cases. The Guardian. URL: https://www.theguardian.com/world/2020/mar/17/israel-to-track-mobile-phones-of-suspected-coronavirus-cases [accessed 2020-11-12]
25. Amit M, Kimhi H, Bader T, Chen J, Glassberg E, Benov A. Mass-surveillance technologies to fight coronavirus spread: the case of Israel. Nat Med 2020 Aug;26(8):1167-1169. [doi: 10.1038/s41591-020-0927-z] [Medline: 32457444]
26. Keshet Y. Fear of panoptic surveillance: using digital technology to control the COVID-19 epidemic. Isr J Health Policy Res 2020 Nov 25;9(1):67 [FREE Full text] [doi: 10.1186/s13584-020-00429-7] [Medline: 33239094]
27. Bergstrøm II. Norway's coronavirus tracing app halted by Data Protection Authority – too invasive and not useful. Sciencenorway. URL: https://sciencenorway.no/covid19-epidemic-society-and-culture/norways-coronavirus-tracing-app-halted-by-data-protection-authority-too-invasive-and-not-useful/1699843 [accessed 2020-11-12]
28. Qatar: contact tracing app security flaw exposed sensitive personal details of more than one million. Amnesty International. URL: https://www.amnesty.org/en/latest/news/2020/05/qatar-covid19-contact-tracing-app-security-flaw/ [accessed 2020-11-12]
29. Covid Tracker App. Health Service Executive. URL: https://www.hse.ie/eng/services/news/newsfeatures/covid19-updates/covid-tracker-app/covid-tracker-app.html [accessed 2021-05-28]
30. O'Brien C. HSE says fix for Covid Tracker app rolled out to all Android users. The Irish Times. URL: https://www.irishtimes.com/business/technology/hse-says-fix-for-covid-tracker-app-rolled-out-to-all-android-users-1.4326646 [accessed 2020-11-26]
31. Active Irish COVID-19 tracing app users drop on battery problem - HSE. Reuters. URL: https://www.reuters.com/article/us-health-coronavirus-ireland-apps-idUSKBN25N1PA [accessed 2020-11-26]
32. Rekanar K, O'Keeffe IR, Buckley S, Abbas M, Beecham S, Chochlov M, et al. Sentiment analysis of user feedback on the HSE's Covid-19 contact tracing app. Ir J Med Sci 2021 Feb 18:1-10 [FREE Full text] [doi: 10.1007/s11845-021-02529-y] [Medline: 33604836]
33. Foxe K. Covid app: 150,000 uninstalled app after August battery issue. Irish Examiner. URL: https://www.irishexaminer.com/news/arid-40058456.html [accessed 2021-01-06]

34. Sidhu R, Appleton G. How the NHS COVID-19 app is making the most of cutting-edge global technology. Technology in the NHS. URL: https://healthtech.blog.gov.uk/2020/10/29/how-the-nhs-covid-19-app-is-making-the-most-of-cutting-edge-global-technology/ [accessed 2020-11-12]
35. Mageit S. NHS COVID-19 contact tracing app fails to ask users to self-isolate. Healthcare IT News. URL: https://www.healthcareitnews.com/news/emea/nhs-covid-19-contact-tracing-app-fails-ask-users-self-isolate [accessed 2020-11-12]
36. Buckley J, Abbas M, Chochlov M, Fitzgerald B, Glynn L, Johnson K, et al. Covigilant: optimizing digital contact tracing from end-user/current practice/idealized-solution perspectives. A proposal for the SFI Covid-19 rapid response funding call. Lero. URL: https://lero.ie/sites/default/files/2020-TR-05_Covigilant%20SFI%20Application%20Tech%20Report.pdf [accessed 2021-01-26]
37. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. Ir J Med Sci 2020 Oct 16:1-25 [FREE Full text] [doi: 10.1007/s11845-020-02389-y] [Medline: 33063226]
38. Storni C, Tsvyatkova D, Richardson I, Buckley J, Abbas M, Beecham S, et al. Toward a compare and contrast framework for COVID-19 contact tracing mobile applications: a look at usability. 2021 Presented at: 14th International Joint Conference on Biomedical Engineering Systems and Technologies; February 11-13, 2021; Vienna, Austria. [doi: 10.5220/0010307005570565]
39. Welsh T, Rekanar K, Abbas M, Chochlov M, Fitzgerald B, Glynn L, et al. Towards a taxonomy for evaluating societal concerns of contact tracing apps. 2020 Presented at: 7th International Conference on Behavioural and Social Computing; November 5-7, 2020; Bournemouth, United Kingdom.
40. Mobile applications in support of contact tracing for COVID-19: a guidance for EU/EEA Member States. European Centre for Disease Prevention and Control. 2020. URL: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-mobile-applications-contact-tracing.pdf [accessed 2021-02-04]
41. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak. European Data Protection Board. 2020. URL: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-042020-use-location-data-and-contact-tracing_en [accessed 2021-01-26]
42. Mobile applications to support contact tracing in the EU's fight against COVID-19: common EU toolbox for Member States. European Commission. 2020. URL: https://ec.europa.eu/health/sites/default/files/ehealth/docs/covid-19_apps_en.pdf [accessed 2021-01-26]
43. Cho H, Ippolito D, Yu YW. Contact tracing mobile apps for COVID-19: privacy considerations and related trade-offs. arXiv. Preprint posted online on March 30, 2020.
44. Bay J, Kek J, Tan A, Hau CS, Yongquan L, Tan J, et al. BlueTrace: a privacy-preserving protocol for community-driven contact tracing across borders. BlueTrace. 2020. URL: https://bluetrace.io/static/bluetrace_whitepaper-938063656596c104632def383eb33b3c.pdf [accessed 2021-01-26]
45. World Health Organization Institutional Repository for Information Sharing. URL: https://apps.who.int/iris/ [accessed 2020-11-26]
46. Publications and data. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/publications-data [accessed 2020-11-26]
47. COVID – 19: Human Rights implications of digital contact tracing technology. Scottish Human Rights Commission. 2020. URL: https://www.scottishhumanrights.com/media/2028/contact-tracing-briefing-180520-final.pdf [accessed 2021-01-26]
48. Upshur REG. Principles for the justification of public health intervention. Can J Public Health 2002;93(2):101-103 [FREE Full text] [Medline: 11968179]
49. Childress JF, Faden RR, Gaare RD, Gostin LO, Kahn J, Bonnie RJ, et al. Public health ethics: mapping the terrain. In: Holland SM, editor. Arguing About Bioethics. New York, NY: Routledge; 2012:361-373.
50. Kass NE. An ethics framework for public health. Am J Public Health 2001 Nov;91(11):1776-1782. [doi: 10.2105/ajph.91.11.1776] [Medline: 11684600]
51. Parker MJ, Fraser C, Abeler-Dörner L, Bonsall D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. J Med Ethics 2020 Jul;46(7):427-431 [FREE Full text] [doi: 10.1136/medethics-2020-106314] [Medline: 32366705]
52. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing: interim guidance, 28 May 2020. World Health Organization. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Ethics_Contact_tracing_apps-2020.1 [accessed 2021-03-17]
53. Ranisch R, Nijsingh N, Ballantyne A, van Bergen A, Buyx A, Friedrich O, et al. Digital contact tracing and exposure notification: ethical guidance for trustworthy pandemic management. Ethics Inf Technol 2020 Oct 21:1-10 [FREE Full text] [doi: 10.1007/s10676-020-09566-8] [Medline: 33106749]
54. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet Digit Health 2020 Aug;2(8):e425-e434 [FREE Full text] [doi: 10.1016/S2589-7500(20)30137-0] [Medline: 32835200]
55. Lo B, Sim I. Ethical framework for assessing manual and digital contact tracing for COVID-19. Ann Intern Med 2021 Mar;174(3):395-400 [FREE Full text] [doi: 10.7326/M20-5834] [Medline: 33076694]

56. Morley J, Cowls J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. Nature 2020 Jun;582(7810):29-31. [doi: 10.1038/d41586-020-01578-0] [Medline: 32467596]
57. Coronavirus: Commission adopts Recommendation to support exit strategies through mobile data and apps. European Commission. 2020 Apr 08. URL: https://ec.europa.eu/commission/presscorner/detail/en/IP_20_626 [accessed 2020-08-23]
58. Ethical considerations in responding to the COVID-19 pandemic. The Nuffield Council on Bioethics. 2020 Mar 17. URL: https://www.nuffieldbioethics.org/publications/ethical-considerations-in-responding-to-the-covid-19-pandemic [accessed 2021-03-17]
59. Sharon T. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. Ethics Inf Technol 2020 Jul 18:1-13. [doi: 10.1007/s10676-020-09547-x] [Medline: 32837287]
60. Braithwaite I, Callender T, Bullock M, Aldridge R. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. Lancet Digit Health 2020 Nov;2(11):e607-e621 [FREE Full text] [doi: 10.1016/S2589-7500(20)30184-9] [Medline: 32839755]
61. Anglemyer A, Moore T, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database Syst Rev 2020 Aug 18;8:CD013699. [doi: 10.1002/14651858.CD013699] [Medline: 33502000]
62. von Wyl V, Bonhoeffer S, Bugnion E, Puhan MA, Salathé M, Stadler T, et al. A research agenda for digital proximity tracing apps. Swiss Med Wkly 2020 Jul 13;150:w20324. [doi: 10.4414/smw.2020.20324] [Medline: 32672340]
63. Steinhauer J, Goodnough A. Contact tracing is failing in many states. Here's why. The New York Times. 2020 Jul 21. URL: https://www.nytimes.com/2020/07/31/health/covid-contact-tracing-tests.html [accessed 2021-01-25]
64. Baumgärtner L, Dmitrienko A, Freisleben B, Gruler A, Höchst J, Kühlberg J, et al. arXiv. Preprint posted online on November 6, 2020.
65. Couch DL, Robinson P, Komesaroff PA. COVID-19-extending surveillance and the panopticon. J Bioeth Inq 2020 Dec;17(4):809-814 [FREE Full text] [doi: 10.1007/s11673-020-10036-5] [Medline: 32840859]
66. Rowe F. Contact tracing apps and values dilemmas: a privacy paradox in a neo-liberal world. Int J Inf Manage 2020 Dec;55:102178 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102178] [Medline: 32836636]
67. Iacobucci G. Government's plan to digitise NHS risks wasting billions, MPs warn. BMJ 2020 Nov 06;371:m4317. [doi: 10.1136/bmj.m4317] [Medline: 33158850]
68. von Wyl V. Challenges for nontechnical implementation of digital proximity tracing during the COVID-19 pandemic: media analysis of the SwissCovid app. JMIR Mhealth Uhealth 2021 Feb 26;9(2):e25345 [FREE Full text] [doi: 10.2196/25345] [Medline: 33606658]
69. Berman G, Carter K, Herranz MG, Sekara V. Digital contact tracing and surveillance during COVID-19. General and child-specific ethical issues. UNICEF Office of Research. 2020. URL: https://www.unicef-irc.org/publications/1096-digital-contact-tracing-surveillance-covid-19-response-child-specific-issues-iwp.html [accessed 2020-08-23]
70. Leith DJ, Farrell S. Measurement-based evaluation of Google/Apple Exposure Notification API for proximity detection in a light-rail tram. PLoS One 2020;15(9):e0239943 [FREE Full text] [doi: 10.1371/journal.pone.0239943] [Medline: 32997724]
71. Lockey S, Edwards M, Hornsey M, Gillespie N, Akhlaghpour S, Colville S. Profiling adopters (and non-adopters) of a contact tracing mobile application: insights from Australia. Int J Med Inform 2021 May;149:104414. [doi: 10.1016/j.ijmedinf.2021.104414] [Medline: 33639328]
72. Williams S, Armitage C, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]
73. Guillon M, Kergall P. Attitudes and opinions on quarantine and support for a contact-tracing application in France during the COVID-19 outbreak. Public Health 2020 Nov;188:21-31 [FREE Full text] [doi: 10.1016/j.puhe.2020.08.026] [Medline: 33059232]
74. Thomas R, Michaleff Z, Greenwood H, Abukmail E, Glasziou P. Concerns and misconceptions about the Australian Government's COVIDSafe App: cross-sectional survey study. JMIR Public Health Surveill 2020 Nov 04;6(4):e23081 [FREE Full text] [doi: 10.2196/23081] [Medline: 33048826]
75. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in Switzerland: panel survey analysis. JMIR Public Health Surveill 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: 10.2196/25701] [Medline: 33326411]
76. Garrett PM, White JP, Lewandowsky S, Kashima Y, Perfors A, Little DR, et al. The acceptability and uptake of smartphone tracking for COVID-19 in Australia. PLoS One 2021;16(1):e0244827 [FREE Full text] [doi: 10.1371/journal.pone.0244827] [Medline: 33481841]
77. Zimmermann B, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. J Med Internet Res 2021 Feb 08;23(2):e25525 [FREE Full text] [doi: 10.2196/25525] [Medline: 33503000]
78. Kaspar K. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: quantitative survey study. J Med Internet Res 2020 Aug 27;22(8):e21613 [FREE Full text] [doi: 10.2196/21613] [Medline: 32759100]

79. Oldeweme A, Märtins J, Westmattelmann D, Schewe G. The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: empirical study on the adoption of COVID-19 tracing apps. J Med Internet Res 2021 Feb 08;23(2):e25893 [FREE Full text] [doi: 10.2196/25893] [Medline: 33465036]
80. Zhang M, Chow A, Smith H. COVID-19 contact-tracing apps: analysis of the readability of privacy policies. J Med Internet Res 2020 Dec 03;22(12):e21572 [FREE Full text] [doi: 10.2196/21572] [Medline: 33170798]
81. Vokinger K, Nittas V, Witt C, Fabrikant S, von Wyl V. Digital health and the COVID-19 epidemic: an assessment framework for apps from an epidemiological and legal perspective. Swiss Med Wkly 2020 May 04;150:w20282. [doi: 10.4414/smw.2020.20282] [Medline: 32418194]
82. Health Service Executive. GitHub. URL: https://github.com/HSEIreland/ [accessed 2021-01-25]
83. Corona-Warn-App. GitHub. URL: https://github.com/corona-warn-app [accessed 2021-01-25]
84. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey evidence. medRxiv. Preprint posted online on May 8, 2020. [doi: 10.1101/2020.05.05.20091587]
85. Hassandoust F, Akhlaghpour S, Johnston A. Individuals' privacy concerns and adoption of contact tracing mobile applications in a pandemic: a situational privacy calculus perspective. J Am Med Inform Assoc 2021 Mar 01;28(3):463-471. [doi: 10.1093/jamia/ocaa240] [Medline: 33164077]
86. Dave R, Gupta R. Mandating the use of proximity tracking apps during coronavirus disease 2019: ethical justifications. Front Med (Lausanne) 2020;7:590265. [doi: 10.3389/fmed.2020.590265] [Medline: 33344477]
87. Ivers LC, Weitzner DJ. Can digital contact tracing make up for lost time? Lancet Public Health 2020 Aug;5(8):e417-e418 [FREE Full text] [doi: 10.1016/S2468-2667(20)30160-2] [Medline: 32682488]
88. Currie D, Peng C, Lyle D, Jameson B, Frommer M. Stemming the flow: how much can the Australian smartphone app help to control COVID-19? Public Health Res Pract 2020 Jun 30;30(2):3022009. [doi: 10.17061/phrp3022009] [Medline: 32601652]
89. Shamil M, Farheen F, Ibtehaz N, Khan I, Rahman M. An agent-based modeling of COVID-19: validation, analysis, and recommendations. Cognit Comput 2021 Feb 19:1-12 [FREE Full text] [doi: 10.1007/s12559-020-09801-w] [Medline: 33643473]
90. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. Lancet Respir Med 2020 Jun;8(6):547-548 [FREE Full text] [doi: 10.1016/S2213-2600(20)30228-9] [Medline: 32401711]
91. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. Lancet Public Health 2020 May;5(5):e256 [FREE Full text] [doi: 10.1016/S2468-2667(20)30061-X] [Medline: 32199471]
92. Kumar D, Hemmige V, Kallen M, Giordano T, Arya M. Mobile phones may not bridge the digital divide: a look at mobile phone literacy in an underserved patient population. Cureus 2019 Feb 20;11(2):e4104 [FREE Full text] [doi: 10.7759/cureus.4104] [Medline: 31057998]
93. Singapore distributes Covid contact-tracing tokens. BBC. URL: https://www.bbc.com/news/business-54143015#:~:text=Singapore%20has%20started%20distributing%20Bluetooth,was%20rolled%20out%20in%20March [accessed 2020-11-13]
94. Anderez DO, Kanjo E, Pogrebna G, Kaiwartya O, Johnson SD, Hunt JA. A COVID-19-based modified epidemiological model and technological approaches to help vulnerable individuals emerge from the lockdown in the UK. Sensors (Basel) 2020 Sep 02;20(17):4967 [FREE Full text] [doi: 10.3390/s20174967] [Medline: 32887338]
95. Wilmink G, Summer I, Marsyla D, Sukhu S, Grote J, Zobel G, et al. Real-time digital contact tracing: development of a system to control COVID-19 outbreaks in nursing homes and long-term care facilities. JMIR Public Health Surveill 2020 Aug 25;6(3):e20828 [FREE Full text] [doi: 10.2196/20828] [Medline: 32745013]
96. Goel S, Nagpal R, Mehrotra D. Mobile applications usability parameters: taking an insight view. In: Mishra DK, Nayak MK, Joshi A, editors. Information and Communication Technology for Sustainable Development. Singapore: Springer; 2018:35-43.
97. Gelman DL. Design For Kids: Digital Products for Playing and Learning. New York, NY: Rosenfeld Media; 2014.
98. Soni N, Aloba A, Morga KS, Wisniewski PJ, Anthony L. A framework of touchscreen interaction design recommendations for children (TIDRC): characterizing the gap between research evidence and design practice. In: Proceedings of the 18th ACM International Conference on Interaction Design and Children. 2019 Jun Presented at: IDC '19; June 2019; Boise, ID p. 419-431. [doi: 10.1145/3311927.3323149]
99. Alonso-Ríos D, Vázquez-García A, Mosqueira-Rey E, Moret-Bonillo V. Usability: a critical analysis and a taxonomy. Int J Hum-Computer Interaction 2010 Jan 06;26(1):53-74. [doi: 10.1080/10447310903025552]
100. Kaur E, Haghighi PD. A context-aware usability model for mobile health applications. In: Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media. 2016 Nov Presented at: MoMM '16; November 2016; Singapore, Singapore p. 181-189. [doi: 10.1145/3007120.3007135]
101. Kasali FA, Taiwo OO, Akinyemi IO, Alaba OB, Awodele O, Kuyoro SO. An enhanced usability model for mobile health application. Int J Computer Sci Inf Secur 2019 Feb;17(2):20-29.

102. Kascak LR, Rébola CB, Sanford JA. Integrating Universal Design (UD) principles and mobile design guidelines to improve design of mobile health applications for older adults. 2014 Presented at: 2014 IEEE International Conference on Healthcare Informatics; September 15-17, 2014; Verona, Italy p. 343-348. [doi: 10.1109/ichi.2014.54]
103. Harrington CN, Ruzic L, Sanford J. Universally accessible mHealth apps for older adults: towards increasing adoption and sustained engagement. In: Antona M, Stephanidis C, editors. Universally accessible mHealth apps for older adults: towards increasing adoption and sustained engagement. Cham: Springer; 2017:3-12.
104. Ballantyne M, Jha A, Jacobsen A, Hawker JS, El-Glaly YN. Study of accessibility guidelines of mobile applications. In: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia. 2018 Nov Presented at: MUM 2018; November 2018; Cairo, Egypt p. 305-315. [doi: 10.1145/3282894.3282921]
105. Horstmann K, Buecker S, Krasko J, Kritzler S, Terwiel S. Who does or does not use the 'Corona-Warn-App' and why? Eur J Public Health 2021 Feb 01;31(1):49-51 [FREE Full text] [doi: 10.1093/eurpub/ckaa239] [Medline: 33340328]
106. Staats R. Designing UI with color blind users in mind. Secret Stache. 2020. URL: https://www.secretstache.com/blog/designing-for-color-blind-users/ [accessed 2021-01-23]
107. Sanwikarja P. Contact tracing: how do you design an app millions of people will trust? UX Collective. 2020. URL: https://uxdesign.cc/how-do-you-design-an-app-millions-of-people-will-trust-8a63f5a5660a [accessed 2021-01-23]
108. COVID-19: governments must promote and protect access to and free flow of information during pandemic – international experts. United Nations Human Rights Office of the High Commissioner. 2020. URL: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=25729&LangID=E [accessed 2021-01-23]
109. Regulation (EU) 2016/679 of the European Parliment and of the Council of 27 April 2016. EUR-Lex. 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN [accessed 2017-09-20]
110. Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies. EUR-Lex. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L2102 [accessed 2021-03-18]
111. Accessibility requirements for ICT products and services. ETSI. URL: https://www.etsi.org/deliver/etsi_en/301500_301599/301549/02.01.02_60/en_301549v020102p.pdf [accessed 2021-01-23]
112. Web Content Accessibility Guidelines (WCAG) overview. World Wide Web Consortium. 2020. URL: https://www.w3.org/WAI/standards-guidelines/wcag/ [accessed 2021-01-23]
113. Xcertia mHealth App Guidelines. Healthcare Information and Management Systems Society. 2019. URL: https://www.himss.org/sites/hde/files/media/file/2020/04/17/xcertia-guidelines-2019-final.pdf [accessed 2021-01-23]
114. Aman. 2020. URL: https://amanapp.jo/en [accessed 2021-01-23]
115. Open-Source Project Corona-Warn-App. 2020. URL: https://www.coronawarn.app/en/ [accessed 2021-01-23]
116. NOVID. 2020. URL: https://www.novid.org/#how-it-works [accessed 2020-11-26]
117. PathCheck Foundation. 2020. URL: https://www.pathcheck.org/ [accessed 2021-01-23]
118. Stop COVID. Gov.pl. 2020. URL: https://www.gov.pl/web/protegosafe [accessed 2021-01-23]
119. Tosi J, Taffoni F, Santacatterina M, Sannino R, Formica D. Performance evaluation of Bluetooth Low Energy: a systematic review. Sensors (Basel) 2017 Dec 13;17(12):2898 [FREE Full text] [doi: 10.3390/s17122898] [Medline: 29236085]
120. Gibney S, McCarthy T. Profile of smartphone ownership and use in Ireland. Gov.ie. 2020. URL: https://assets.gov.ie/81401/e6c10ac5-e6b4-438e-b31d-30cd2b241f7c.pdf [accessed 2020-08-10]
121. Lewandowsky S, Dennis S, Perfors A, Kashima Y, White JP, Garrett P, et al. Public acceptance of privacy-encroaching policies to address the COVID-19 pandemic in the United Kingdom. PLoS One 2021;16(1):e0245740 [FREE Full text] [doi: 10.1371/journal.pone.0245740] [Medline: 33481877]
122. Bachtiger P, Adamson A, Quint JK, Peters NS. Belief of having had unconfirmed Covid-19 infection reduces willingness to participate in app-based contact tracing. NPJ Digit Med 2020 Nov 06;3(1):146. [doi: 10.1038/s41746-020-00357-5] [Medline: 33299071]
123. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mölken M. COVID-19 contact tracing apps: predicted uptake in the Netherlands based on a discrete choice experiment. JMIR Mhealth Uhealth 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: 10.2196/20741] [Medline: 32795998]
124. Garrett P, Wang Y, White J, Hsieh S, Strong C, Lee YC, et al. Young adults view smartphone tracking technologies for COVID-19 as acceptable: the case of Taiwan. Int J Environ Res Public Health 2021 Feb 02;18(3):1332 [FREE Full text] [doi: 10.3390/ijerph18031332] [Medline: 33540628]
125. Cavoukian A. Privacy by design: the 7 foundational principles. Information and Privacy Commissioner of Ontario. URL: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf [accessed 2021-03-18]
126. Lauristin M. Draft Report on the proposal for a regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications). European Parliament. 2017 Sep. URL: https://www.europarl.europa.eu/doceo/document/LIBE-PR-606011_EN.pdf?redirect [accessed 2020-08-23]
127. Guidance note: legal bases for processing personal data. Data Protection Commission. 2019 Dec. URL: https://www.dataprotection.ie/sites/default/files/uploads/2020-04/Guidance%20on%20Legal%20Bases.pdf [accessed 2021-03-18]

128. Principles of data protection. Data Protection Commission. URL: https://www.dataprotection.ie/en/individuals/data-protection-basics/principles-data-protection [accessed 2020-09-26]
129. Gibney S, Bruton L, Doherty P. COVID contact tracing app: user perspectives and experience research. Gov.ie. 2020. URL: https://assets.gov.ie/81402/f470c4e5-d188-4bbd-8aa1-27af31b07f7f.pdf [accessed 2020-08-23]
130. Loh PS. Accuracy of Bluetooth-Ultrasound contact tracing: experimental results from NOVID iOS Version 2.1 using five-year-old phones. NOVID. 2020. URL: https://www.novid.org/downloads/20200625-accuracy.pdf [accessed 2020-08-23]
131. Leith DJ, Farrell S. GAEN due diligence: verifying the Google/Apple Covid Exposure Notification API. School of Computer Science and Statistics Trinity College Dublin. 2021. URL: https://www.scss.tcd.ie/Doug.Leith/pubs/gaen_verification.pdf [accessed 2020-08-23]
132. Leith DJ, Farrell S. Measurement-based evaluation of Google/Apple Exposure Notification API for proximity detection in a commuter bus. arXiv. Preprint posted online on June 15, 2020. [doi: 10.1371/journal.pone.0250826] [Medline: 33914810]
133. Bengio Y, Ippolito D, Janda R, Jarvie M, Prud'homme B, Rousseau JF, et al. Inherent privacy limitations of decentralized contact tracing apps. J Am Med Inform Assoc 2021 Jan 15;28(1):193-195 [FREE Full text] [doi: 10.1093/jamia/ocaa153] [Medline: 32584990]
134. Nakamoto I, Jiang M, Zhang J, Zhuang W, Guo Y, Jin M, et al. Evaluation of the design and implementation of a peer-to-peer COVID-19 contact tracing mobile app (COCOA) in Japan. JMIR Mhealth Uhealth 2020 Dec 01;8(12):e22098 [FREE Full text] [doi: 10.2196/22098] [Medline: 33170801]
135. Lewis D. Contact-tracing apps help reduce COVID infections, data suggest. Nature 2021 Mar;591(7848):18-19. [doi: 10.1038/d41586-021-00451-y] [Medline: 33623147]
136. Patil S, Lu H, Saunders C, Potoglou D, Robinson N. Public preferences for electronic health data storage, access, and sharing - evidence from a pan-European survey. J Am Med Inform Assoc 2016 Nov;23(6):1096-1106 [FREE Full text] [doi: 10.1093/jamia/ocw012] [Medline: 27107445]
137. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
138. Contact transmission of COVID-19 in South Korea: novel investigation techniques for tracing contacts. Osong Public Health Res Perspect 2020 Feb;11(1):60-63. [doi: 10.24171/j.phrp.2020.11.1.09] [Medline: 32149043]
139. Park S, Choi GJ, Ko H. Information technology-based tracing strategy in response to COVID-19 in South Korea-privacy controversies. JAMA 2020 Jun 02;323(21):2129-2130. [doi: 10.1001/jama.2020.6602] [Medline: 32324202]
140. Valentino-DeVries J, Singer N, Keller MH, Krolik A. Your apps know where you were last night, and they're not keeping it secret. The New York Times. URL: https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html [accessed 2021-01-13]
141. RENEW EUROPE Webinar on COVID-19 contact tracing applications. Renew Europe Group: Livecasts. URL: https://re.livecasts.eu/webinar-on-contact-tracing-applications/ [accessed 2020-09-30]
142. Barth S, de Jong MDT. The privacy paradox – investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review. Telematics Inform 2017 Nov;34(7):1038-1058. [doi: 10.1016/j.tele.2017.04.013]
143. Nguyen CT, Saputra YM, Van Huynh N, Nguyen NT, Khoa TV, Tuan BM, et al. Enabling and emerging technologies for social distancing: a comprehensive survey and open problems. arXiv. Preprint posted online on September 22, 2020.
144. Simko L, Chang JL, Jiang M, Calo R, Roesner F, Kohno T. COVID-19 contact tracing and privacy: a longitudinal study of public opinion. arXiv. Preprint posted online on December 4, 2020.
145. Scherr TF, DeSousa JM, Moore CP, Hardcastle A, Wright DW. App use and usability of a barcode-based digital platform to augment COVID-19 contact tracing: postpilot survey and paradata analysis. JMIR Public Health Surveill 2021 Mar 26;7(3):e25859 [FREE Full text] [doi: 10.2196/25859] [Medline: 33630745]
146. Huang Z, Guo H, Lee Y, Ho E, Ang H, Chow A. Performance of digital contact tracing tools for COVID-19 response in Singapore: cross-sectional study. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e23148 [FREE Full text] [doi: 10.2196/23148] [Medline: 33006944]
147. Meyer S, Windisch T, Witt N, Dzibela D. Google Exposure Notification API testing fraunhofer IIS. GitHub. 2020. URL: https://github.com/corona-warn-app/cwa-documentation/blob/master/2020_06_24_Corona_API_measurements.pdf [accessed 2021-01-28]
148. Liu S, Striegel A. Accurate extraction of face-to-face proximity using smartphones and Bluetooth. In: Proceedings of 20th International Conference on Computer Communications and Networks. 2011 Presented at: 2011 ICCCN; July 31-August 4, 2011; Lahaina, HI. [doi: 10.1109/icccn.2011.6006081]
149. Stanley J, Granick J. The limits of location tracking in an epidemic. American Civil Liberties Union. 2020. URL: https://www.aclu.org/sites/default/files/field_document/limits_of_location_tracking_in_an_epidemic.pdf [accessed 2021-02-04]
150. Salathé M, Althaus C, Anderegg N, Antonioli D, Ballouz T, Bugnon E, et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. Swiss Med Wkly 2020 Dec 14;150:w20457. [doi: 10.4414/smw.2020.20457] [Medline: 33327003]

151. Rodríguez P, Graña S, Alvarez-León EE, Battaglini M, Darias FJ, Hernán MA, RadarCovidPilot Group, et al. A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. Nat Commun 2021 Jan 26;12(1):587. [doi: 10.1038/s41467-020-20817-6] [Medline: 33500407]
152. Issue guide: COVID-19 case investigation and contact tracing: considerations for using digital technologies. Association of State and Territorial Health Officials. 2020. URL: https://www.astho.org/ASTHOReports/COVID-19-Case-Investigation-and-Contact-Tracing-Considerations-for-Using-Digital-Technologies/07-16-20/ [accessed 2021-02-04]
153. Loh PS. Flipping the perspective in contact tracing. arXiv. Preprint posted online on November 13, 2020.
154. A rapid solution to help combat the impact of Covid-19. NearForm. URL: https://www.nearform.com/services/contact-tracing-app-development/ [accessed 2020-09-30]
155. Troncoso C, Payer M, Hubaux JP, Salathé M, Larus J, Bugnion E, et al. Decentralized privacy-preserving proximity tracing. GitHub. 2020 May 25. URL: https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf [accessed 2020-09-30]
156. Pan-European Privacy-Preserving Proximity Tracing. URL: https://www.pepp-pt.org/ [accessed 2020-09-30]
157. Zhang B, Kreps S, McMurry N, McCain RM. Americans' perceptions of privacy and surveillance in the COVID-19 pandemic. PLoS One 2020;15(12):e0242652 [FREE Full text] [doi: 10.1371/journal.pone.0242652] [Medline: 33362218]
158. Close contact. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/appendix.html#contact [accessed 2020-11-03]
159. Contact tracing: public health management of persons, including healthcare workers, having had contact with COVID-19 cases in the European Union – second update. European Centre for Disease Prevention and Control. 2020. URL: https://tinyurl.com/5apdf98z [accessed 2020-08-10]
160. Qureshi Z, Jones N, Temple R, Larwood JPJ, Greenhalgh T, Bourouiba L. What is the evidence to support the 2-metre social distancing rule to reduce COVID-19 transmission? The Centre for Evidence-Based Medicine. 2020. URL: https://www.cebm.net/covid-19/what-is-the-evidence-to-support-the-2-metre-social-distancing-rule-to-reduce-covid-19-transmission/#:~:text=Social%20distancing%20rules%20are%20based,be%20sufficient%20to%20reduce%20risk [accessed 2020-08-10]
161. Epidemiological motivation of the transmission risk level. GitHub. 2020 Oct. URL: https://raw.githubusercontent.com/corona-warn-app/cwa-documentation/master/transmission_risk.pdf [accessed 2021-02-04]
162. Yasaka T, Lehrich B, Sahyouni R. Peer-to-peer contact tracing: development of a privacy-preserving smartphone app. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
163. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
164. Saw Y, Tan E, Liu J, Liu J. Predicting public uptake of digital contact tracing during the COVID-19 pandemic: results from a nationwide survey in Singapore. J Med Internet Res 2021 Feb 03;23(2):e24730 [FREE Full text] [doi: 10.2196/24730] [Medline: 33465034]
165. Jehi L, Ji X, Milinovich A, Erzurum S, Rubin BP, Gordon S, et al. Individualizing risk prediction for positive coronavirus disease 2019 testing: results from 11,672 patients. Chest 2020 Oct;158(4):1364-1375 [FREE Full text] [doi: 10.1016/j.chest.2020.05.580] [Medline: 32533957]
166. Almagor J, Picascia S. Exploring the effectiveness of a COVID-19 contact tracing app using an agent-based model. Sci Rep 2020 Dec 17;10(1):22235. [doi: 10.1038/s41598-020-79000-y] [Medline: 33335125]
167. Menges D, Aschmann H, Moser A, Althaus CL, von Wyl V. The role of the SwissCovid digital proximity tracing app during the pandemic response: results for the Canton of Zurich. medRxiv. Preprint posted online on February 3, 2021. [doi: 10.1101/2021.02.01.21250972]
168. Ballouz T, Menges D, Aschmann HE, Domenghino A, Fehr JS, Puhan MA, et al. Digital proximity tracing app notifications lead to faster quarantine in non-household contacts: results from the Zurich SARS-CoV-2 Cohort Study. medRxiv. Preprint posted online on December 23, 2020. [doi: 10.1101/2020.12.21.20248619]
169. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 app. GitHub. URL: https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Epidemiological_Impact_of_the_NHS_COVID_19_App_Public_Release_V1.pdf [accessed 2021-03-12]
170. Masel J, Shilen A, Helming B, Rutschman J, Windham G, Judd M, et al. Quantifying meaningful adoption of a SARS-CoV-2 exposure notification app on the campus of the University of Arizona. medRxiv. Preprint posted online on February 3, 2021. [doi: 10.1101/2021.02.02.21251022]
171. Lin L. How China slowed coronavirus: lockdowns, surveillance, enforcers. The Wall Street Journal. URL: https://www.wsj.com/articles/how-china-slowed-coronavirus-lockdowns-surveillance-enforcers-11583868093 [accessed 2020-12-03]
172. Roderick P, Macfarlane A, Pollock AM. Getting back on track: control of covid-19 outbreaks in the community. BMJ 2020 Jun 25;369:m2484. [doi: 10.1136/bmj.m2484] [Medline: 32586845]
173. Data Protection Impact Assessment: COVID Tracker app. GitHub. URL: https://github.com/HSEIreland/covidtracker-documentation/blob/master/documentation/privacy/Data%20Protection%20Impact%20Assessment%20for%20the%20COVID%20Tracker%20App%20-%2026.06.2020.pdf [accessed 2020-08-10]

174. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946. World Health Organization. URL: http://apps.who.int/iris/bitstream/handle/10665/44192/9789241650472_eng.pdf?sequence=1 [accessed 2020-08-10]
175. Human Rights Quarterly. The Siracusa Principles on the limitation and derogation provisions in the International Covenant on Civil and Political Rights. Hum Rights Q 1985 Feb;7(1):3. [doi: 10.2307/762035]
176. Storeng K, de Bengy Puyvallée A. The smartphone pandemic: how big tech and public health authorities partner in the digital response to Covid-19. Glob Public Health 2021 Feb 18:1-17. [doi: 10.1080/17441692.2021.1882530] [Medline: 33602063]
177. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733 [FREE Full text] [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
178. Concern over UK government's "ad hoc system" for covid-19 tracking, testing, and contact tracing. The BMJ. 2020. URL: https://www.bmj.com/company/newsroom/concern-over-uk-governments-ad-hoc-system-for-covid-19-tracking-testing-and-contact-tracing/ [accessed 2020-08-10]
179. How does the Corona warning app determine an increased risk? GitHub. URL: https://github.com/corona-warn-app/cwa-documentation/blob/master/translations/cwa-risk-assessment.de.md [accessed 2020-08-10]
180. Mobile accessibility: how WCAG 2.0 and other W3C/WAI guidelines apply to mobile. World Wide Web Consortium. 2015. URL: https://www.w3.org/TR/mobile-accessibility-mapping/ [accessed 2021-01-23]
181. Art. 8 GDPR: conditions applicable to child's consent in relation to information society services. General Data Protection Regulation. 2018. URL: https://gdpr-info.eu/art-8-gdpr/ [accessed 2020-10-01]
182. COVID-19 contact tracing: data protection expectations on app development. Information Commissioner's Office. 2020. URL: https://ico.org.uk/media/for-organisations/documents/2617676/ico-contact-tracing-recommendations.pdf [accessed 2020-06-15]
183. Alarifi A, Al-Salman A, Alsaleh M, Alnafessah A, Al-Hadhrami S, Al-Ammar M, et al. Ultra wideband indoor positioning technologies: analysis and recent advances. Sensors (Basel) 2016 May 16;16(5):707 [FREE Full text] [doi: 10.3390/s16050707] [Medline: 27196906]
184. Stone M. What is ultra-wideband, and how does it work? Samsung Business Insights. URL: https://insights.samsung.com/2020/08/21/what-is-ultra-wideband-and-how-does-it-work/ [accessed 2020-09-26]
185. Department of Health and the HSE announce the publication of the Covid Tracker App Data Protection Impact Assessment and source code. Gov.ie. URL: https://www.gov.ie/en/press-release/bb5d9-department-of-health-and-the-hse-today-announce-the-publication-of-the-covid-tracker-app-data-protection-impact-assessment-and-source-code/ [accessed 2020-11-04]
186. Convention for the protection of human rights and fundamental freedoms. European Court of Human Rights. URL: https://www.echr.coe.int/documents/convention_eng.pdf [accessed 2020-09-30]
187. Consolidated text: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). EUR-Lex. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02016R0679-20160504 [accessed 2020-11-04]
188. Fox C. Google hit with £44m GDPR fine over ads. BBC. URL: https://www.bbc.com/news/technology-46944696 [accessed 2020-11-04]
189. Google accused by rival of fundamental GDPR breaches. Financial Times. URL: https://www.ft.com/content/66dbc3ba-848a-4206-8b97-27c0e384ff27 [accessed 2020-11-04]
190. Gorey C. The Covid Tracker Ireland app has updated, so what's new? Silicon Republic. URL: https://www.siliconrepublic.com/enterprise/the-covid-tracker-ireland-app-update-battery-issue [accessed 2020-11-04]
191. van Diggelen F, Enge P. The world's first GPS MOOC and worldwide lab using smartphones. Stanford University. 2015. URL: https://web.stanford.edu/group/scpnt/pnt/PNT15/2015_Presentation_Files/I15-vanDiggelen-GPS_MOOC-Smartphones.pdf [accessed 2020-08-10]
192. Hernández-Orallo E, Calafate CT, Cano J, Manzoni P. Evaluating the Effectiveness of COVID-19 Bluetooth-Based Smartphone Contact Tracing Applications. Applied Sciences 2020 Oct 13;10(20):7113. [doi: 10.3390/app10207113]
193. Colizza V, Grill E, Mikolajczyk R, Cattuto C, Kucharski A, Riley S, et al. Time to evaluate COVID-19 contact-tracing apps. Nat Med 2021 Mar;27(3):361-362. [doi: 10.1038/s41591-021-01236-6] [Medline: 33589822]
194. O'Callaghan ME, Abbas M, Buckley J, Fitzgerald B, Johnson K, Laffey J, et al. A national survey on use and perceived usefulness of the Irish COVID tracker contact tracing app. JMIR Preprints. Preprint posted online on March 25, 2021. [doi: 10.2196/preprints.28960]
195. Arifeen MM, Al Mamun A, Kaiser MS, Mahmud M. Blockchain-enable contact tracing for preserving user privacy during Covid-19 outbreak. Preprints. Preprint posted online on July 22, 2020. [doi: 10.20944/preprints202007.0502.v1]
196. Idrees S, Nowostawski M, Jameel R. Blockchain-based digital contact tracing apps for COVID-19 pandemic management: issues, challenges, solutions, and future directions. JMIR Med Inform 2021 Feb 09;9(2):e25245 [FREE Full text] [doi: 10.2196/25245] [Medline: 33400677]

197. Shubina V, Holcer S, Gould M, Lohan ES. Survey of decentralized solutions with mobile devices for user location tracking, proximity detection, and contact tracing in the COVID-19 era. Data 2020 Sep 23;5(4):87. [doi: 10.3390/data5040087]
198. How Apple's ultra-wideband chip could transform its products. Financial Times. 2019 Sep 20. URL: https://www.ft.com/content/47e914a0-da3b-11e9-8f9b-77216ebe1f17 [accessed 2021-03-15]
199. Hsu J. Can AI make Bluetooth contact tracing better? IEEE Spectrum. URL: https://spectrum.ieee.org/the-human-os/artificial-intelligence/machine-learning/ai-bluetooth-contact-tracing [accessed 2021-03-17]

Abbreviations

API: application programming interface
DCTA: digital contact tracing app
DP-3T: Decentralized Privacy Preserving Proximity Tracing
ECDC: European Centre for Disease Prevention and Control
EDPB: European Data Protection Board
GDPR: General Data Protection Regulation
IDCTA: ideal digital contact tracing app
LE: Low Energy
NHS: National Health Service
PEPP-PT: Pan-European Privacy Preserving Proximity Tracing
ROI: Republic of Ireland
WHO: World Health Organization

Edited by L Buis; submitted 04.02.21; peer-reviewed by H Mehdizadeh, T Alessa; comments to author 26.02.21; revised version received 17.03.21; accepted 05.04.21; published 07.06.21.

Please cite as:
O'Connell J, Abbas M, Beecham S, Buckley J, Chochlov M, Fitzgerald B, Glynn L, Johnson K, Laffey J, McNicholas B, Nuseibeh B, O'Callaghan M, O'Keeffe I, Razzaq A, Rekanar K, Richardson I, Simpkin A, Storni C, Tsvyatkova D, Walsh J, Welsh T, O'Keeffe D
Best Practice Guidance for Digital Contact Tracing Apps: A Cross-disciplinary Review of the Literature
JMIR Mhealth Uhealth 2021;9(6):e27753
URL: https://mhealth.jmir.org/2021/6/e27753
doi:10.2196/27753
PMID:34003764

Review

Effects of Telemedicine and mHealth on Systolic Blood Pressure Management in Stroke Patients: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Meina Lv[1*], BS; Tingting Wu[1*], BS; Shaojun Jiang[1], BS; Wenjun Chen[1], BS; Jinhua Zhang[1], PhD

Department of Pharmacy, Fujian Medical University Union Hospital, Fujian Medical University, Fuzhou, China

[*]these authors contributed equally

Corresponding Author:
Jinhua Zhang, PhD
Department of Pharmacy
Fujian Medical University Union Hospital
Fujian Medical University
29 Xinquan Road
Fuzhou, 350001
China
Phone: 86 591 83357896 ext 8301
Fax: 86 591 83357896
Email: pollyzhang2006@126.com

Abstract

Background: Stroke is a common, harmful disease with high recurrence and mortality rates. Uncontrolled blood pressure is an important and changeable risk factor for stroke recurrence. Telemedicine and mobile health (mHealth) interventions may have the potential to facilitate the control of blood pressure among stroke survivors, but their effect has not been established.

Objective: This systematic review and meta-analysis of randomized controlled trials (RCTs) was conducted to estimate the effects of telemedicine and mHealth interventions on the control of systolic blood pressure among stroke survivors.

Methods: The research literature published up to June 28, 2020, and consisting of RCTs related to telemedicine and mHealth interventions was searched in PubMed, EMBASE, Web of Science, and the Cochrane Library. The Cochrane risk of bias tool (RoB 2.0) was used to evaluate the quality of the studies. The Cochran *Q* test and I^2 statistic were used to assess heterogeneity. Data were meta-analyzed using a random-effects model. Mean difference (MD) with 95% CI and 95% prediction interval (PI) were calculated.

Results: In total, 9 RCTs with a total sample size of 1583 stroke survivors met the inclusion criteria. Compared with the usual care, telemedicine and mHealth had a significantly greater impact on the control of systolic blood pressure (MD –5.49; 95% CI –7.87 to –3.10; *P*<.001; 95% PI –10.46 to –0.51). A subgroup analysis showed that the intervention mode of telephone plus SMS text messaging (MD –9.09; 95% CI –12.71 to –5.46; *P*<.001) or only telephone (MD –4.34; 95% CI –6.55 to –2.13; *P*<.001; 95% PI –7.24 to –1.45) had a greater impact on the control of systolic blood pressure than usual care. Among the stroke survivors with an intervention interval ≤1 week (MD –6.51; 95% CI –9.36 to –3.66; *P*<.001; 95% PI –12.91 to –0.10) or a baseline systolic blood pressure ≥140 mm Hg (MD –6.15; 95% CI –9.44 to –2.86; *P*<.001; 95% PI –13.55 to 1.26), the control of systolic blood pressure using telemedicine and mHealth was better than that of usual care.

Conclusions: In general, telemedicine and mHealth reduced the systolic blood pressure of stroke survivors by an average of 5.49 mm Hg compared with usual care. Telemedicine and mHealth are a relatively new intervention mode with potential applications for the control of systolic blood pressure among stroke survivors, especially those with hypertensive stroke.

(JMIR Mhealth Uhealth 2021;9(6):e24116) doi:10.2196/24116

KEYWORDS

stroke; systolic blood pressure; mHealth; telemedicine; meta-analysis; self-management

Introduction

Stroke is a common, harmful disease and a main cause of death and disability worldwide. It has the characteristics of high morbidity, disability, recurrence, and mortality [1-3]. Stroke survivors have a high risk of recurrence, with recurrent stroke entailing more severe symptoms and worse results than the first occurrence [4,5]. Stroke not only affects patients' quality of life but also imposes an economic burden on the family, medical system, and society [6,7]. However, about 85% of stroke cases are preventable, and effective secondary prevention can reduce the recurrence rate of stroke [8-10].

Noncommunicable diseases are the main causes for the increase in the incidence of stroke. Approximately 90.5% of global stroke diseases can be attributed to modifiable risk factors, among which hypertension is the most common for first and recurrent strokes but is modifiable [11-13]. Uncontrolled blood pressure is an important changeable risk factor for stroke recurrence. Implementing secondary preventive measures can reduce the recurrence of stroke by 80% [14]. A recent systematic review and meta-regression analysis emphasized that strict and active blood pressure control may be the most critical treatment strategy for the secondary prevention of stroke, highlighting that the reduction in systolic blood pressure is linearly related to reduction in the risk of recurrent cerebrovascular events [15]. However, many stroke survivors have the risk factor of high blood pressure [14]. More than one-third of patients continue to have poor blood pressure control following a stroke or transient ischemic attack, but most people are unaware of these risks [14,16].

Therefore, some researchers have tried to reduce the risk of recurrent cerebrovascular events through interventions to improve blood pressure after a stroke or transient ischemic attack. Telemedicine and mobile health (mHealth) interventions have a potential role in this endeavor. An increasing number of studies have been conducted on the use of telemedicine and mHealth interventions to manage systolic blood pressure in stroke survivors [17-25], but it is not clear whether their effect is better than that of usual care. To objectively evaluate the efficacy of these interventions and provide a reference for clinical application, this study adopted the Cochrane evaluation method to conduct a systematic review and meta-analysis of existing international randomized controlled trials (RCTs) related to telemedicine and mHealth for control of systolic blood pressure in stroke survivors.

Methods

Data Sources and Search Strategy

This study follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [26]. We conducted a comprehensive literature search in online databases, including PubMed, EMBASE, Web of Science, and the Cochrane Library. In order to conduct a comprehensive search, we also searched Chinese literature, gray literature, and the reference lists of the studies yielded by the original search. We searched relevant studies published until June 28, 2020. The search keywords were as follows: "stroke" OR "brain infarction" OR "transient ischemic attack" OR "cerebral hemorrhage" OR "subarachnoid hemorrhage," "mobile applications" OR "telemedicine" OR "text messaging" OR "cell phone" OR "smartphone" OR "social media" OR "internet," and "blood pressure" OR "hypertension." A detailed search strategy for each database is presented in Multimedia Appendix 1. The literature search and screening were carried out independently by 2 researchers (ML and TW).

Inclusion Criteria

We included all studies that met the following requirements: the study's design was an RCT, participants were diagnosed with a stroke (hemorrhagic stroke or ischemic stroke) or transient ischemic attack, interventions were provided for patients using telemedicine (with telemedicine defined as the provision of health services at a distance using a range of technologies, such as telephone, telemonitoring, etc [27,28]) and mHealth (with mHealth defined as the delivery of health service through mobile and wireless applications, including mobile phones, SMS text messaging, wearable devices, etc [29]), the control group received usual care, and the main outcome indicator was systolic blood pressure.

Exclusion Criteria

Studies were excluded from the meta-analysis if any of the telemedicine and mHealth intervention or usual care management was independently discussed, or if the original research data were incomplete or unusable and useful data could not be obtained by contacting the original author.

Data Extraction

The data were retrieved from the selected studies. The extracted data included study information (author, publication year, country), study characteristics (study population, sample size), participants' characteristics (age, gender, baseline systolic blood pressure), intervention information (intervention mode, intervention interval), and main outcome indicators (systolic blood pressure). The required data were extracted independently by 2 researchers (ML and TW) and cross-referenced to avoid potential extraction errors. All disagreements were discussed with a third researcher to reach a consensus.

Quality Assessment

Two independent researchers used the Cochrane risk of bias tool (RoB 2.0) [30] to evaluate the quality of the selected literature. The items addressed were as follows: bias arising from the randomization process, deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, bias in the selection of the reported result, and overall risk of bias. An additional researcher was asked to conduct an evaluation to help resolve disputes that arose during the evaluation process.

Statistical Analysis

Stata version 14.0 (StataCorp) was used for the meta-analysis. The Cochran Q test and I^2 statistic were used to assess heterogeneity [31]. In the heterogeneity assessment, I^2 is considered to be nil if it is below 25%, low if it is 25%-50%, moderate if it is 51%-75%, and high if it is above 75% [32].

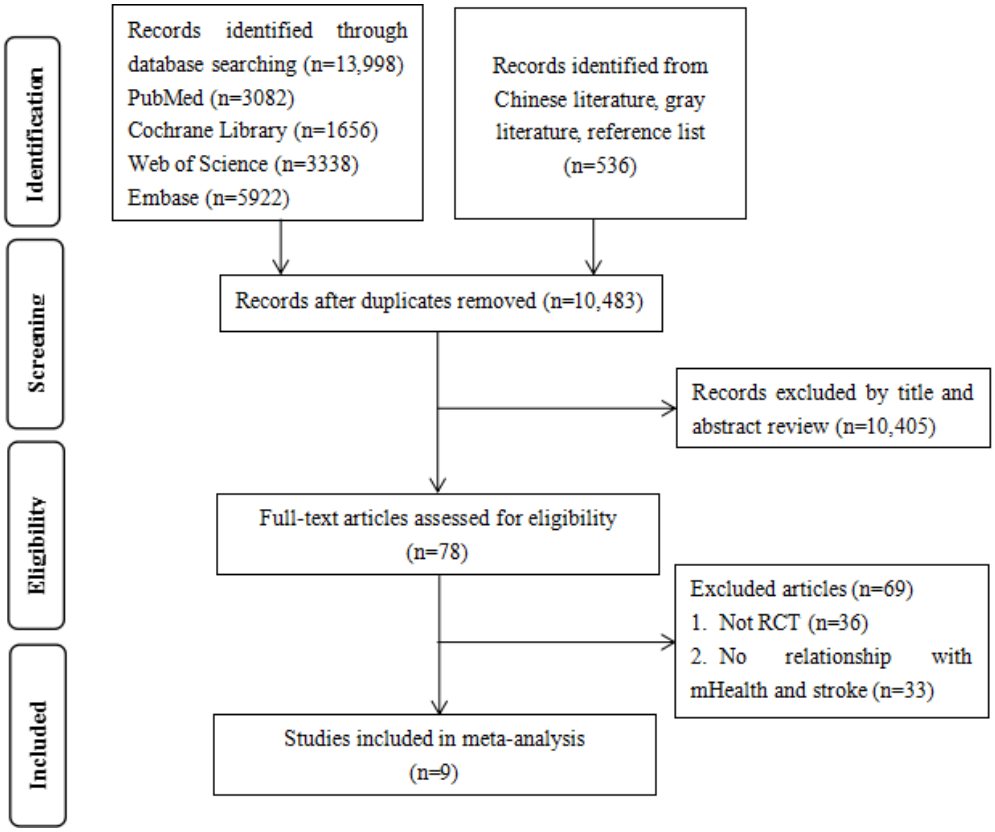
Due to expected heterogeneity (study characteristics and the manner in which studies were conducted) between studies, a random-effects model was used to estimate the mean difference (MD) with 95% CI being considered the statistic of interest [33]. In addition, the 95% prediction interval (PI) was calculated for the overall weighted mean estimate [34]. To explore the factors influencing mHealth interventions, we conducted subgroup analyses of the intervention mode, intervention interval, and baseline systolic blood pressure. Interrater agreement was calculated by using the κ statistic according to the following scheme: κ value <0, worse than that expected by chance; 0.21-0.40, poor; 0.41-0.60, moderate; 0.61-0.80, good; and 0.81-1.00, very good level of agreement [35]. Publication bias was evaluated by inspection of funnel plots and Egger tests [36]. In this study, a *P* value <.05 was considered statistically significant.

Results

Study Selection

A total of 13,998 studies were retrieved using the search strategy. After screening, 9 studies were included in the meta-analysis, comprising a total of 1583 patients: 798 in the mHealth intervention group and 785 in the usual care group. The literature screening process and results are shown in Figure 1.

Figure 1. Flow diagram of the selection of studies. mHealth: mobile health; RCT: randomized controlled trial.



Study Characteristics and Quality Assessment

The basic characteristics of the included studies are presented in Multimedia Appendix 2. All of the included studies were RCTs and published in 2010 or later, which is in line with the rapid development and spread of mHealth technology in recent years. These studies were conducted in different countries and regions, of which 3 were from the United Kingdom [17-19], 2 from the United States [20,21], 2 from China [23,24], 1 from Ghana [22], and 1 from Sweden [25]. The participants were stroke survivors. The mean or median age of the patients ranged from 54.3 years to 73.5 years. The proportion of women ranged from 23.1% to 60.0%. We used the Cochrane risk of bias tool (RoB 2.0) to evaluate the risk of bias in the 9 included studies. The results showed that risk of bias was deemed to be either "low" or "with some concerns" (Table 1). A κ value of 0.768 (95% CI 0.673-0.841; *P*<.001) in this study indicated that there was a good agreement between encoders.

Table 1. Cochrane risk-of-bias tool for randomized controlled trials (RoB 2.0).

Study	Bias arising from the randomization process	Deviations from intended interventions	Bias due to missing outcome data	Bias in measurement of the outcome	Bias in selection of the reported result	Overall risk of bias
Adie et al (2010) [17]	Low	Low	Low	Low	Low	Low[a]
Hanley et al (2015) [18]	Some concerns	Some concerns	Low	Low	Low	Some concerns[b]
Kerry et al (2013) [19]	Low	Low	Low	Low	Low	Low
Lakshminarayan et al (2018) [20]	Low	Some concerns	Low	Low	Low	Some concerns
Mackenzie et al (2013) [21]	Low	Some concerns	Low	Low	Some concerns	Some concerns
Sarfo et al (2018) [22]	Low	Low	Low	Low	Low	Low
Wan et al (2018) [23]	Low	Low	Low	Low	Low	Low
Wang et al (2020) [24]	Low	Low	Low	Low	Low	Low
Ögren et al (2018) [25]	Low	Some concerns	Some concerns	Low	Low	Some concerns

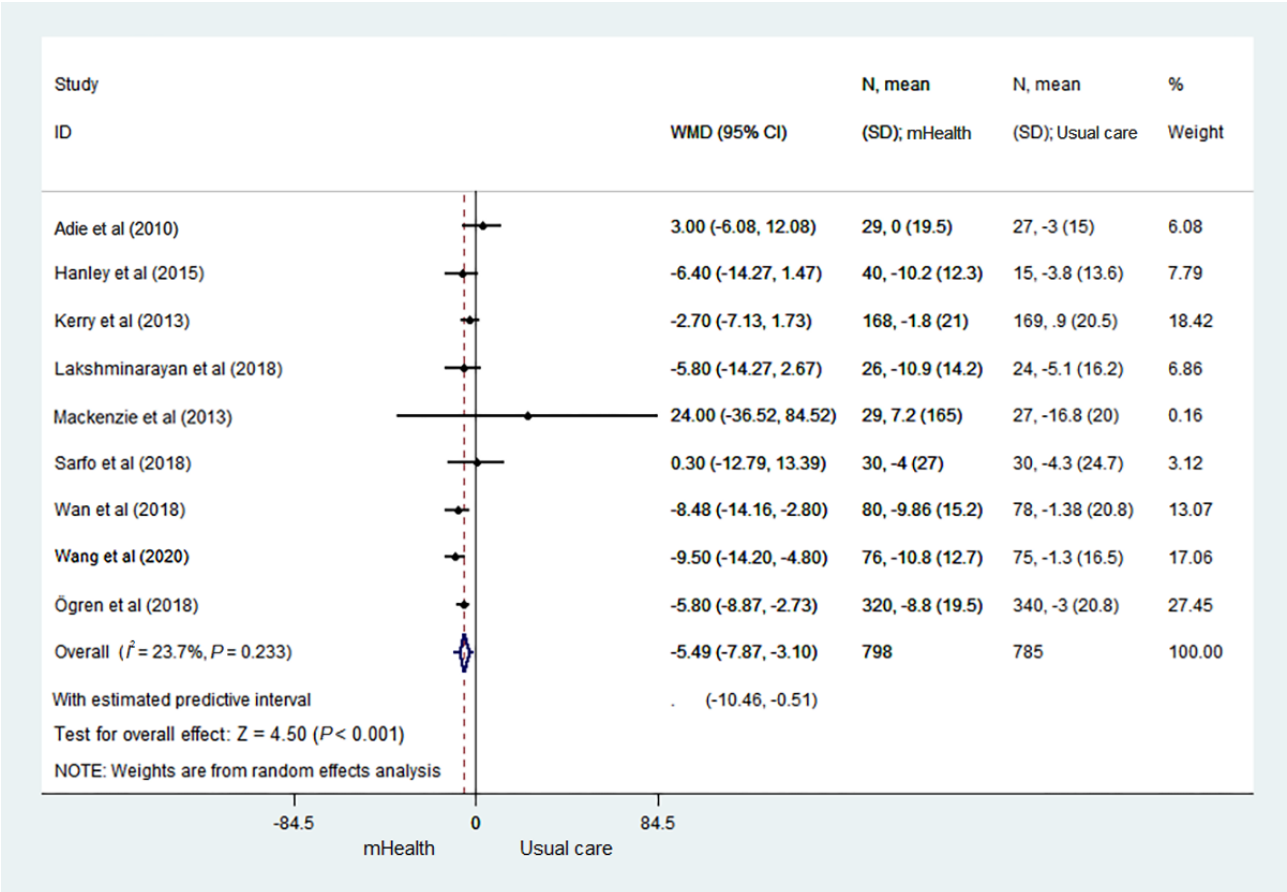
[a]Low: when present in this column, this indicates the study is judged to be at low risk of bias for all domains for this result.

[b]Some concerns: when present in this column, the study is judged to raise some concerns in at least one domain for this result, but not due to a high risk of bias for any domain.

Comparison of Changes in Systolic Blood Pressure

Figure 2 [17-25] illustrates the changes in systolic blood pressure between the 2 groups in the 9 studies. There was statistical heterogeneity between studies (I^2=23.7%). The results showed that the control of systolic blood pressure of the stroke survivors in the telemedicine and mHealth group was better than that of the stroke survivors in the usual care group, and the difference was statistically significant (MD –5.49; 95% CI –7.87 to –3.10; P<.001; 95% PI –10.46 to –0.51).

Figure 2. Forest plot of the systolic blood pressure of the telemedicine and mHealth group and usual care group. mHealth: mobile health; WMD: weighted mean difference.



Subgroup Analyses

We conducted subgroup analyses of the intervention mode. When the intervention mode consisted of telephone plus SMS text messaging or only telephone, the telemedicine and mHealth group showed a larger effect on the control of systolic blood pressure than did the usual care group, with an MD of –9.09 (95% CI –12.71 to –5.46; *P*<.001) and –4.34 (95% CI –6.55 to –2.13; *P*<.001; 95% PI –7.24 to –1.45), respectively (Figure 3 [17-25]).

We also performed subgroup analyses of the intervention interval and baseline systolic blood pressure (Figures 4 and 5 [17-25]). Compared to the usual care group, the telemedicine and mHealth group had better control of systolic blood pressure, with an intervention interval ≤1 week, and the difference was statistically significant (MD –6.51; 95% CI –9.36 to –3.66; *P*<.001; 95% PI –12.91 to –0.10). When the intervention interval was greater than 1 week, no significant difference was found in the control of systolic blood pressure between the 2 groups (MD –2.08; 95% CI –10.12 to 5.95; *P*=.61; 95% PI –83.93 to 79.76). In addition, among the stroke survivors with a baseline systolic blood pressure <140 mm Hg, no significant difference in the control of systolic blood pressure was found between the mHealth intervention group and the usual care group (MD –4.04; 95% CI –8.75 to 0.67; *P*=.09; 95% PI –50.34 to 42.25). In contrast, among the stroke survivors with a baseline systolic blood pressure ≥140 mm Hg, the control of systolic blood pressure of the telemedicine and mHealth group was significantly better than that of the usual care group, and the difference was statistically significant (MD –6.15; 95% CI –9.44 to –2.86; *P*<.001; 95% PI –13.55 to 1.26).

Figure 3. Forest plot of the subgroup analysis of the mode of intervention. mHealth: mobile health; WMD: weighted mean difference.

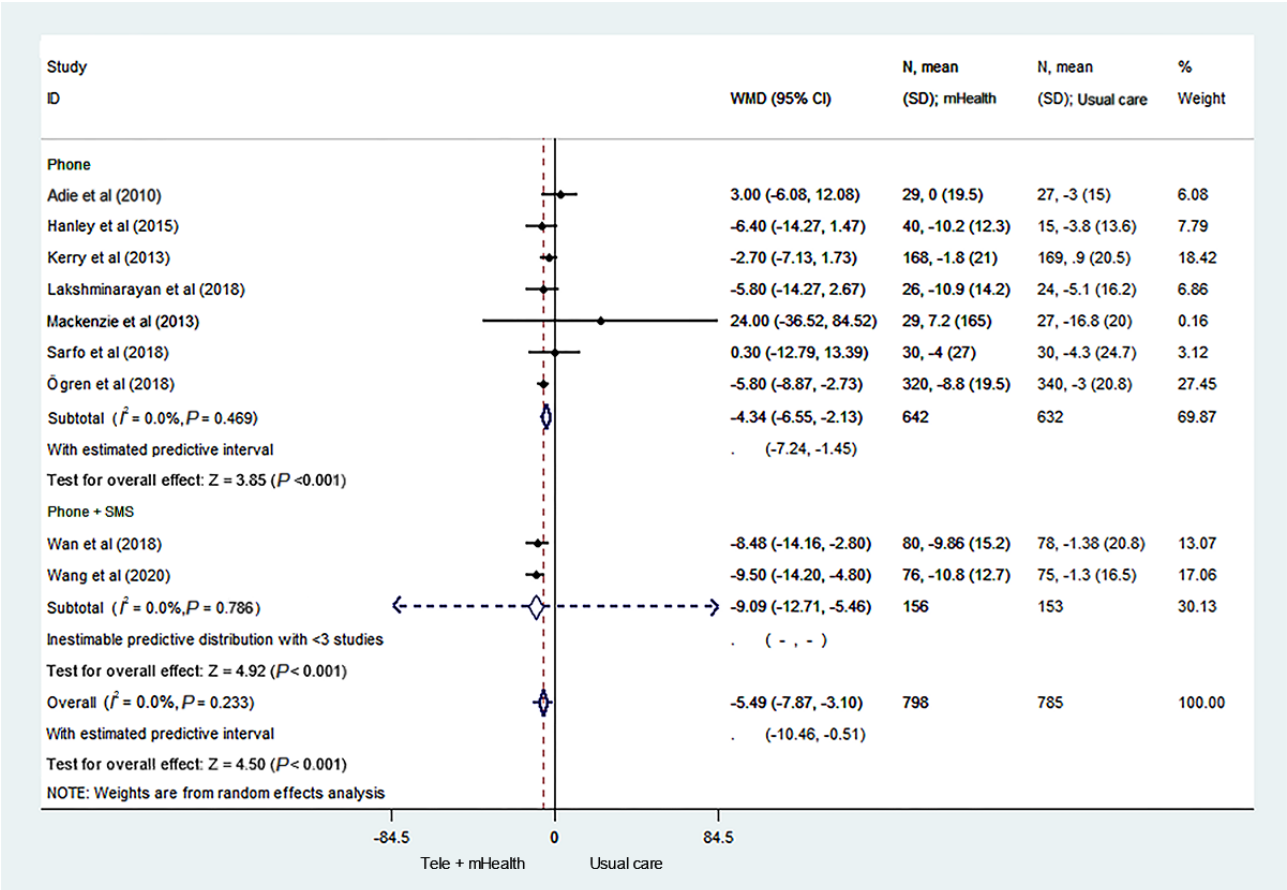


Figure 4. Forest plot of the subgroup analysis of the intervention interval. mHealth: mobile health; WMD: weighted mean difference.

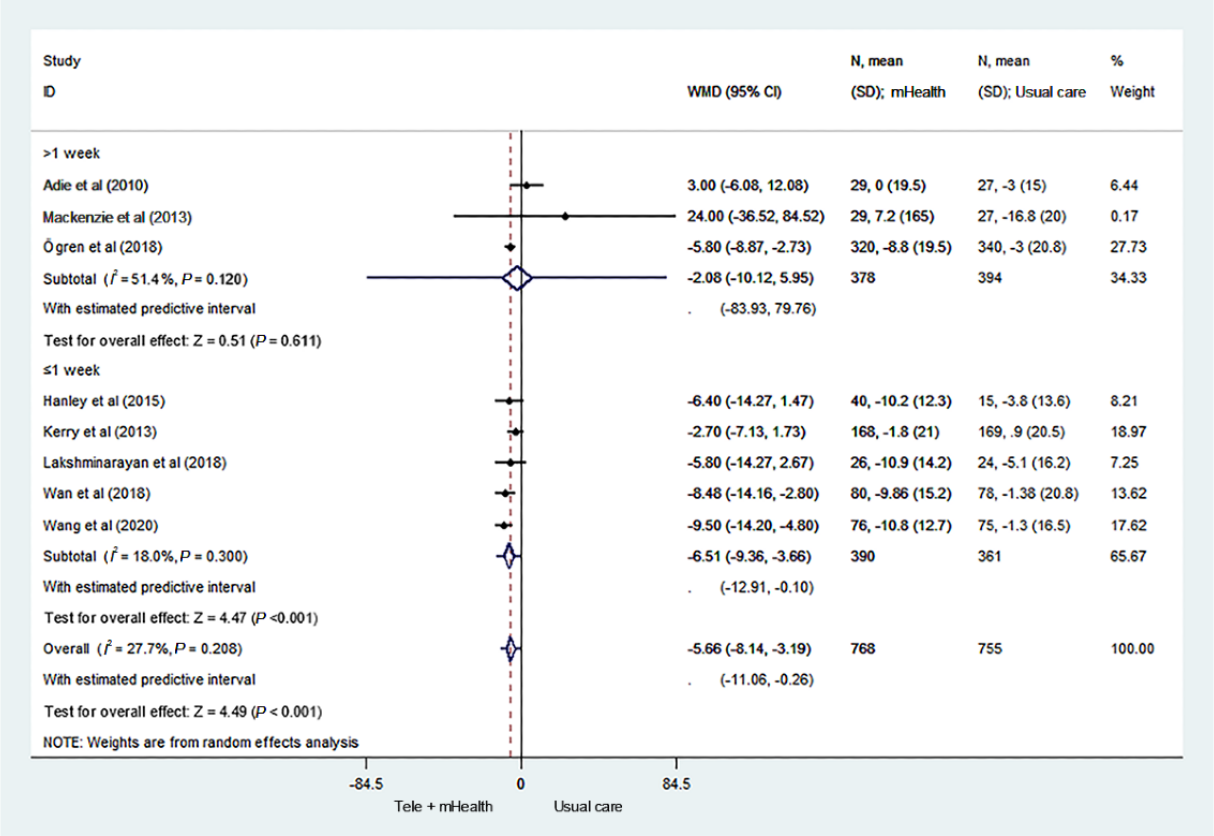
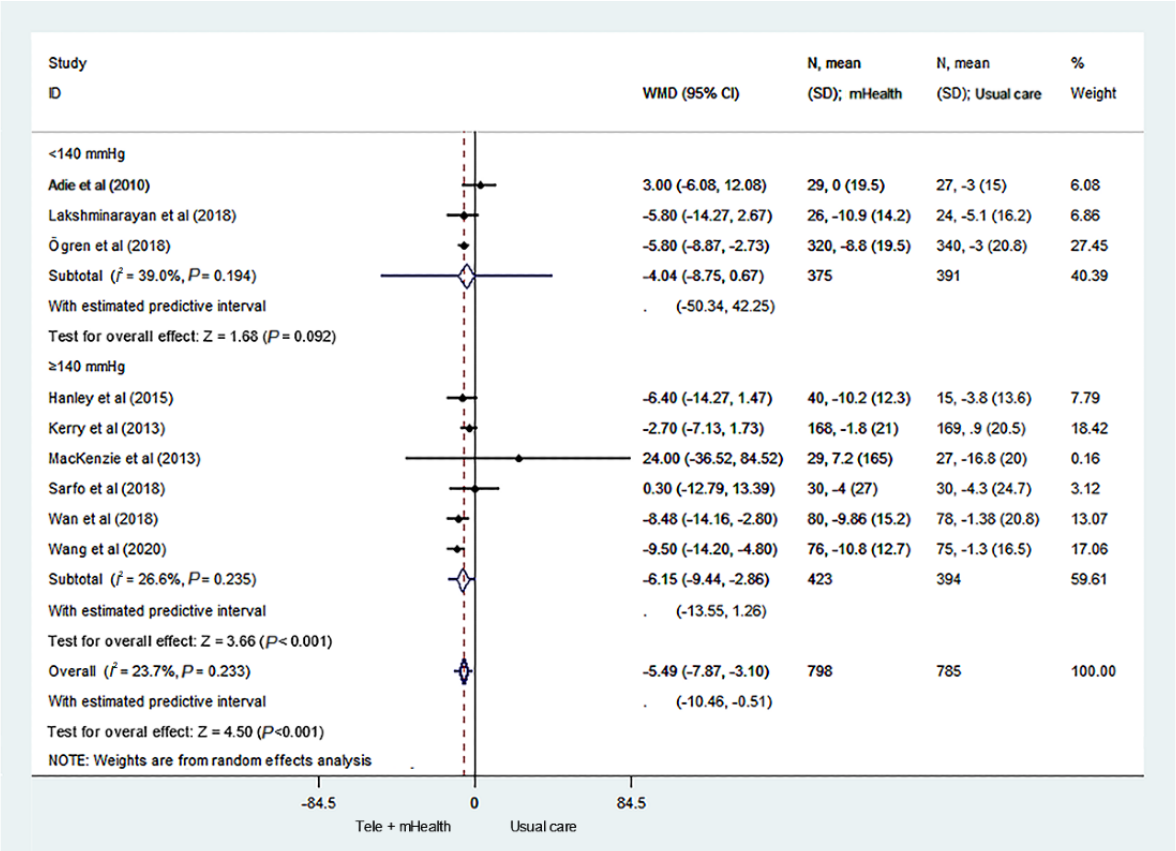


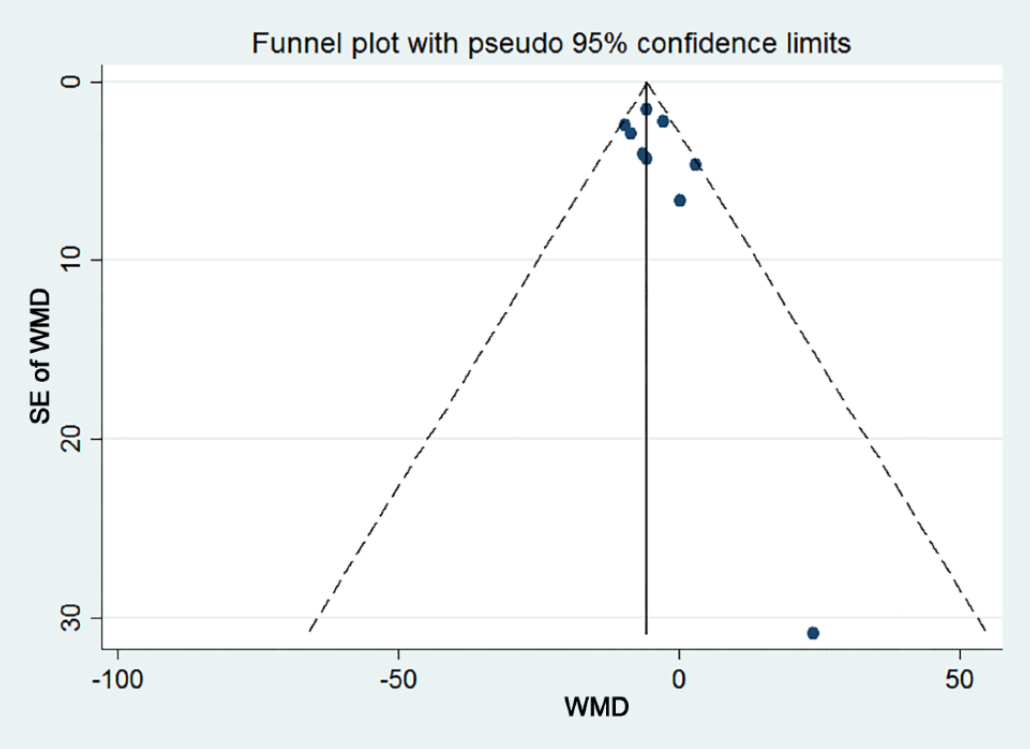
Figure 5. Forest plot of the subgroup analysis of the baseline systolic blood pressures. mHealth: mobile health; WMD: weighted mean difference.



Publication Bias and Sensitivity Analysis

Funnel plot inspection and the Egger test showed no publication bias (*P*=.16; Figure 6). Furthermore, a sensitivity analysis of the outcome indicators of systolic blood pressure was conducted using the method of excluding relevant studies one by one. The results did not change significantly, indicating that the findings of this analysis were stable.

Figure 6. Funnel plot of the systolic blood pressures. WMD: weighted mean difference.



Discussion

Stroke is characterized by high recurrence and mortality rates [1-3]. Hypertension is an important risk factor for stroke recurrence [13], so it is essential for stroke survivors to control their blood pressure. The main purpose of this meta-analysis was to evaluate the effect of telemedicine and mHealth interventions on the control of systolic blood pressure in stroke survivors. We conducted a systematic review and meta-analysis of 9 RCTs. Compared with usual care, the telemedicine and mHealth intervention reduced the systolic blood pressure by an average of 5.49 mm Hg. It is worth mentioning that this change in systolic blood pressure was equivalent to the decrease in the systolic blood pressure reported in a meta-analysis of interventions to improve the lifestyles of patients (eg, reasonable diet, aerobic exercise, restriction of alcohol and sodium intake) [37,38]. Studies have shown that a 3-mm Hg reduction in systolic blood pressure can reduce stroke mortality by 8% [39]. Therefore, telemedicine and mHealth interventions for stroke survivors may be a measure worth considering.

As far as we know, this is the first systematic, quantitative analysis and summary of all available evidence of telemedicine and mHealth interventions for the management of systolic blood pressure in the population of stroke survivors. More importantly, we found that the stroke survivors in the telemedicine and mHealth group had better control of their systolic blood pressure than did the usual care group after receiving interventions that actively sent electronic messages (telephone calls, SMS text messages). This result may be expected because poor self-management and poor compliance are major problems affecting patients' blood pressure control [40]. Stroke survivors with low compliance may benefit more from active interventions, such as telephone calls or SMS text messages. Therefore, active telemedicine and mHealth interventions may yield clinical benefits for stroke survivors by helping them achieve blood pressure control.

This meta-analysis included RCTs from different countries (the United States, the United Kingdom, Sweden, Ghana, and China), indicating that mHealth interventions may be applicable to people in different countries and different medical systems. Furthermore, the average baseline systolic blood pressure was 128.0-154.0 mm Hg, which indicates that the included studies targeted stroke survivors extensively for telemedicine and mHealth interventions. We found that among the stroke survivors with a baseline systolic blood pressure <140 mm Hg, there was no significant difference between the telemedicine and mHealth group and the usual care group. However, for stroke survivors with a baseline systolic blood pressure ≥140 mm Hg, the telemedicine and mHealth group had significantly better control of systolic blood pressure than did the usual care group. This is a major finding in stroke survivors with a baseline systolic blood pressure ≥140 mm Hg, which indicates that telemedicine and mHealth interventions may have greater benefits for stroke survivors with hypertension. If the proper intervention is conducted for an extended period, this may have a significant clinical impact.

Telemedicine and mHealth interventions are becoming an increasingly common way to support patients with chronic diseases in adhering to their medications and conducting self-management [41]. Telemedicine and mHealth interventions can provide reminder strategies and help patients achieve self-monitoring of blood pressure to improve their medical and behavioral management. Nursing staff can make personalized recommendations for blood pressure management based on patients' feedback. We found that when the intervention interval was ≤1 week, the influence on the control of systolic blood pressure of the telemedicine and mHealth group was significantly greater than that of the usual care group. However, there was no significant difference between the 2 groups when the intervention interval was more than 1 week. These findings show that when implementing telemedicine and mHealth interventions for patients, the time interval should be at least 1 week in order to achieve a clinically meaningful effect on the control of systolic blood pressure.

Our research has several limitations worth discussing. First, one of the main limitations is that the duration of the interventions included in the selected studies was relatively short. There was only 1 study over 12 months, and a lack of data from studies lasting more than 12 months makes it impossible to conduct subgroup analyses. Blood pressure control in stroke survivors may be a long-term process, requiring continuous lifestyle changes. It is important to understand the long-term (over 12 months) effectiveness and safety of telemedicine and mHealth interventions in stroke survivors. Thus, more research is needed for further analyses and verification. Second, compared with usual care, the telemedicine and mHealth intervention reduced the systolic blood pressure. Statistically speaking, the difference was significant, but its clinical significance still needs to be confirmed by further study. Third, as most of the included studies only provided limited information on the profiles of the participants, it was impossible to analyze the effects of some factors on the telemedicine and mHealth interventions, such as participants' socioeconomic and educational status and combination of drugs, which still need to be explored further in future research.

Preliminary analysis shows that the telemedicine and mHealth interventions reduced the systolic blood pressure of stroke survivors by 5.49 mm Hg on average compared with patients who received usual care. Telemedicine and mHealth interventions may be an important strategy to promote the control of systolic blood pressure in stroke survivors, and this benefit may be even greater for patients with hypertensive stroke. We also found that telemedicine and mHealth interventions with active reminders via telephone calls or SMS text messages and an intervention interval ≤1 week may be more effective. In short, telemedicine and mHealth interventions are relatively new. If used correctly, they have potential application in the control of systolic blood pressure in stroke survivors, specifically those with hypertensive stroke.

Acknowledgments

This study was funded by the Natural Science Foundation of Fujian Province of China (grant 2018Y0037) and the Fujian Medical Innovation Project of China (grant 2019-CX-19).

Authors' Contributions

JZ initiated the study. ML and TW performed the data extraction and analyses. ML drafted the first version of the manuscript. JZ and ML critically reviewed the manuscript and revised it. All authors made a substantial contribution to the concept and design of the study, interpreted the data, and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[PDF File (Adobe PDF File), 72 KB - mhealth_v9i6e24116_app1.pdf]

Multimedia Appendix 2
Characteristics of the 9 randomized controlled trials included in the study.
[PDF File (Adobe PDF File), 67 KB - mhealth_v9i6e24116_app2.pdf]

References

1. Feigin VL, Mensah GA, Norrving B, Murray CJ, Roth GA. Atlas of the Global Burden of Stroke (1990-2013): The GBD 2013 Study. Neuroepidemiology 2015:230-236 [FREE Full text] [doi: 10.1159/000441106]
2. Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. Lancet Neurology 2007 Feb;6(2):182-187 [FREE Full text] [doi: 10.1016/s1474-4422(07)70031-5]
3. Feigin VL. Stroke in developing countries: can the epidemic be stopped and outcomes improved? Lancet Neurology 2007 Feb;6(2):94-97 [FREE Full text] [doi: 10.1016/s1474-4422(07)70007-8]

4. Leistner S, Michelson G, Laumeier I, Ahmadi M, Smyth M, Nieweler G. Intensified secondary prevention intending a reduction of recurrent events in TIA and minor stroke patients (INSPiRE-TMS): a protocol for a randomised controlled trial. BMC Neurol 2013 Jan 24;13(1):1-8 [FREE Full text] [doi: 10.1186/1471-2377-13-11]
5. Zhang X, Pan J, Wan L, Liu Z, Mo M, Wang M. Factors influencing health behaviour, blood pressure control, and disability in hypertensive ischaemic stroke patients after a comprehensive reminder intervention. J Adv Nurs 2020;76(6):1384-1393 [FREE Full text] [doi: 10.1111/jan.14340]
6. Xu G, Liu X, Wu W, Zhang R, Yin Q. Recurrence after ischemic stroke in Chinese patients: impact of uncontrolled modifiable risk factors. Cerebrovasc Dis 2007;23(2-3):117-120 [FREE Full text] [doi: 10.1159/000097047]
7. Em S, Bozkurt M, Caglayan M, Ceylan Cevik F, Kaya C, Oktayoglu P. Psychological health of caregivers and association with functional status of stroke patients. Topics in Stroke Rehabilitation 2017;24(5):323-329 [FREE Full text] [doi: 10.1080/10749357.2017.1280901]
8. Sarikaya H, Ferro J, Arnold M. Stroke prevention - medical and lifestyle measures. Eur Neurol 2015;73(3-4):150-157 [FREE Full text] [doi: 10.1159/000367652]
9. Jeon MY, Jeong H. Effects of a stroke primary prevention program on risk factors for at-home elderly. Med Sci Monit 2015;21:3696-3703 [FREE Full text] [doi: 10.12659/msm.895519]
10. Wang M, Shen M, Wan L, Mo M, Wu Z, Li L. Effects of a comprehensive reminder system based on the health belief model for patients who have had a stroke on health behaviors, blood pressure, disability, and recurrence from baseline to 6 months. The Journal of Cardiovascular Nursing 2020;35(2):156-164 [FREE Full text] [doi: 10.1097/jcn.0000000000000631]
11. Kernan WN, Ovbiagele B, Black HR, Bravata DM, Chimowitz MI, Ezekowitz MD, American Heart Association Stroke Council, Council on CardiovascularStroke Nursing, Council on Clinical Cardiology,Council on Peripheral Vascular Disease. Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2014 Jul;45(7):2160-2236. [doi: 10.1161/STR.0000000000000024] [Medline: 24788967]
12. O'Donnell MJ, Chin SL, Rangarajan S, Xavier D, Liu L, Zhang H, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. The Lancet 2016 Aug;388(10046):761-775 [FREE Full text] [doi: 10.1016/s0140-6736(16)30506-2]
13. Feigin VL, Roth GA, Naghavi M, Parmar P, Krishnamurthi R, Chugh S, et al. Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet Neurology 2016 Aug;15(9):913-924 [FREE Full text] [doi: 10.1016/s1474-4422(16)30073-4]
14. Doogue R, McCann D, Fitzgerald N, Murphy AW, Glynn LG, Hayes P. Blood pressure control in patients with a previous stroke/transient ischaemic attack in primary care in Ireland: a cross sectional study. BMC Fam Pract 2020 Jul 10;21(1):139-148 [FREE Full text] [doi: 10.1186/s12875-020-01211-z]
15. Katsanos AH, Filippatou A, Manios E, Deftereos S, Parissis J, Frogoudaki A, et al. Blood pressure reduction and secondary stroke prevention. Hypertension 2017 Jan;69(1):171-179 [FREE Full text] [doi: 10.1161/hypertensionaha.116.08485]
16. Alloubani A, Saleh A, Abdelhafiz I. Hypertension and diabetes mellitus as a predictive risk factors for stroke. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 2018 Jul;12(4):577-584 [FREE Full text] [doi: 10.1016/j.dsx.2018.03.009]
17. Adie K, James MA. Does telephone follow-up improve blood pressure after minor stroke or TIA? Age and Ageing 2010 Jul 27;39(5):598-603 [FREE Full text] [doi: 10.1093/ageing/afq085]
18. Hanley J, Fairbrother P, Krishan A, McCloughan L, Padfield P, Paterson M, et al. Mixed methods feasibility study for a trial of blood pressure telemonitoring for people who have had stroke/transient ischaemic attack (TIA). Trials 2015 Mar 25;16(1):117-126 [FREE Full text] [doi: 10.1186/s13063-015-0628-y]
19. Kerry SM, Markus HS, Khong TK, Cloud GC, Tulloch J, Coster D, et al. Home blood pressure monitoring with nurse-led telephone support among patients with hypertension and a history of stroke: a community-based randomized controlled trial. Canadian Medical Association Journal 2012 Nov 05;185(1):23-31 [FREE Full text] [doi: 10.1503/cmaj.120832]
20. Lakshminarayan K, Westberg S, Northuis C, Fuller CC, Ikramuddin F, Ezzeddine M, et al. A mHealth-based care model for improving hypertension control in stroke survivors: Pilot RCT. Contemporary Clinical Trials 2018 Jul;70:24-34 [FREE Full text] [doi: 10.1016/j.cct.2018.05.005]
21. Mackenzie G, Ireland S, Moore S, Heinz I, Johnson R, Oczkowski W. Tailored interventions to improve hypertension management after stroke or TIA--phase II (TIMS II). Can J Neurosci Nurs(1) 2013;35:27-34 [FREE Full text] [Medline: 23687780]
22. Sarfo F, Treiber F, Gebregziabher M, Adamu S, Patel S, Nichols M, et al. PINGS (Phone-Based Intervention Under Nurse Guidance After Stroke). Stroke 2018 Jan;49(1):236-239 [FREE Full text] [doi: 10.1161/strokeaha.117.019591]
23. Wan L, Zhang X, You L, Ruan H, Chen S. The efficacy of a comprehensive reminder system to improve health behaviors and blood pressure control in hypertensive ischemic stroke patients. The Journal of Cardiovascular Nursing 2018;33(6):509-517 [FREE Full text] [doi: 10.1097/jcn.0000000000000496]
24. Wang M, Shen M, Wan L, Mo M, Wu Z, Li L, et al. Effects of a comprehensive reminder system based on the health belief model for patients who have had a stroke on health behaviors, blood pressure, disability, and recurrence from baseline to 6 months. The Journal of Cardiovascular Nursing 2020;35(2):156-164 [FREE Full text] [doi: 10.1097/jcn.0000000000000631]

25. Ögren J, Irewall A, Söderström L, Mooe T. Long-term, telephone-based follow-up after stroke and TIA improves risk factors: 36-month results from the randomized controlled NAILED stroke risk factor trial. BMC Neurol 2018 Sep 21;18(1):153-161 [FREE Full text] [doi: 10.1186/s12883-018-1158-5]
26. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339:b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
27. The Knowledge Network. URL: http://www.knowledge.scot.nhs.uk/telehealthcare.aspx [accessed 2016-10-27]
28. Ming W, Mackillop LH, Farmer AJ, Loerup L, Bartlett K, Levy JC, et al. Telemedicine technologies for diabetes in pregnancy: a systematic review and meta-analysis. J Med Internet Res 2016 Nov 09;18(11):e290 [FREE Full text] [doi: 10.2196/jmir.6556] [Medline: 27829574]
29. Borrelli B, Ritterband LM. Special issue on eHealth and mHealth: Challenges and future directions for assessment, treatment, and dissemination. Health Psychol 2015 Dec;34S:1205-1208. [doi: 10.1037/hea0000323] [Medline: 26651461]
30. 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] [Medline: 31462531]
31. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions version 5.1. The Cochrane Collaboration. URL: http://handbook-5-1.cochrane.org/ [accessed 2011-03-01]
32. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
33. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010 Apr;1(2):97-111. [doi: 10.1002/jrsm.12] [Medline: 26061376]
34. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester: John Wiley & Sons; 2009.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
36. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
37. Dickinson HO, Mason JM, Nicolson DJ, Campbell F, Beyer FR, Cook JV, et al. Lifestyle interventions to reduce raised blood pressure: a systematic review of randomized controlled trials. J Hypertens 2006 Feb;24(2):215-233. [doi: 10.1097/01.hjh.0000199800.72563.26] [Medline: 16508562]
38. Whelton SP, Chin A, Xin X, He J. Effect of aerobic exercise on blood pressure: a meta-analysis of randomized, controlled trials. Ann Intern Med 2002 Apr 02;136(7):493-503. [doi: 10.7326/0003-4819-136-7-200204020-00006] [Medline: 11926784]
39. Stamler J, Rose G, Stamler R, Elliott P, Dyer A, Marmot M. INTERSALT study findings. Public health and medical care implications. Hypertension 1989 Nov;14(5):570-577 [FREE Full text] [doi: 10.1161/01.hyp.14.5.570]
40. Amer M, Ur Rehman N, Nazir S, Ehsan-Ul-Haq M, Sultana M, Jabeen M. Devoid knowledge and low adherence precipitate uncontrolled blood pressure: a depiction from cross-sectional study. Altern Ther Health Med 2020 Jul 15;26(12). [Medline: 32663188]
41. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. Patient Prefer Adherence 2018;12:813-822 [FREE Full text] [doi: 10.2147/PPA.S145647] [Medline: 29785096]

Abbreviations

MD: mean difference
mHealth: mobile health
PI: prediction interval
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by L Buis; submitted 04.09.20; peer-reviewed by Xiang, R Bajpai, KM Kuo; comments to author 21.12.20; revised version received 27.01.21; accepted 11.03.21; published 11.06.21.

Please cite as:
Lv M, Wu T, Jiang S, Chen W, Zhang J
Effects of Telemedicine and mHealth on Systolic Blood Pressure Management in Stroke Patients: Systematic Review and Meta-Analysis of Randomized Controlled Trials
JMIR Mhealth Uhealth 2021;9(6):e24116
URL: https://mhealth.jmir.org/2021/6/e24116
doi:10.2196/24116
PMID:34114961

Review

Barriers to and Facilitators for Using Nutrition Apps: Systematic Review and Conceptual Framework

Laura Maria König[1,2*], BSc, MSc, PhD; Christiane Attig[3*], BSc, MSc; Thomas Franke[4], PhD; Britta Renner[1], PhD

[1]Department of Psychology, University of Konstanz, Konstanz, Germany

[2]Faculty of Life Sciences, University of Bayreuth, Kulmbach, Germany

[3]Department of Psychology, Chemnitz University of Technology, Chemnitz, Germany

[4]Institute for Multimedia and Interactive Systems, University of Lübeck, Lübeck, Germany

[*]these authors contributed equally

Corresponding Author:
Laura Maria König, BSc, MSc, PhD
Department of Psychology
University of Konstanz
Universitätsstraße 10
Konstanz, 78464
Germany
Phone: 49 7531885317
Email: laura.koenig@uni-konstanz.de

Abstract

Background: Nutrition apps are effective in changing eating behavior and diet-related health risk factors. However, while they may curb growing overweight and obesity rates, widespread adoption is yet to be achieved. Hence, profound knowledge regarding factors motivating and hindering (long-term) nutrition app use is crucial for developing design guidelines aimed at supporting uptake and prolonged use of nutrition apps.

Objective: In this systematic review, we synthesized the literature on barriers to and facilitators for nutrition app use across disciplines including empirical qualitative and quantitative studies with current users, ex-users, and nonusers of nutrition apps.

Methods: A systematic literature search including 6 databases (PubMed, Web of Science, PsychINFO, PSYNDEX, PsycArticles, and SPORTDiscus) as well as backward and forward citation search was conducted. Search strategy, inclusion and exclusion criteria, and the planned data extraction process were preregistered. All empirical qualitative and quantitative studies published in German or English were eligible for inclusion if they examined adolescents (aged 13-18) or adults who were either current users, ex-users, and nonusers of nutrition apps. Based on qualitative content analysis, extracted individual barriers and facilitators were grouped into categories.

Results: A total of 28 publications were identified as eligible. A framework with a 3-level hierarchy was designed which grouped 328 individual barriers and facilitators into 23 subcategories, 12 categories, and 4 clusters that focus on either the individual user (goal setting and goal striving, motivation, routines, lack of awareness of knowledge), different aspects of the app and the smartphone (features, usability of the app or food database, technical issues, data security, accuracy/trustworthiness, costs), positive and negative outcomes of nutrition app use, or interactions between the user and their social environment.

Conclusions: The resulting conceptual framework underlines a pronounced diversity of reasons for (not) using nutrition apps, indicating that there is no “one-size-fits-all” approach for uptake and prolonged use of nutrition apps. Hence, tailoring nutrition apps to needs of specific user groups seems promising for increasing engagement.

(JMIR Mhealth Uhealth 2021;9(6):e20037) doi:10.2196/20037

KEYWORDS

nutrition apps; mHealth; digital health; usage facilitators; usage barriers

Introduction

Overweight is one of today's most urgent public health issues [1,2]. It is related to a number of noncommunicable diseases including cardiovascular diseases, cancer, diabetes, premature deaths, and reduced quality of life [3-5]. Furthermore, it poses substantial economic costs on society [6]. Currently, more than 2 billion people worldwide are affected and figures are projected to rise further [2,7]. Thus, interventions are urgently needed to curb growing overweight and obesity rates by promoting weight loss, weight maintenance, or preventing weight gain [2]. Ideally, these interventions are not only effective, but also affordable and reach a large number of people around the globe.

The widespread adoption of mobile phones [8] thus constitutes a promising opportunity to reduce overweight. Accordingly, using mobile phones and other mobile devices for health purposes (mobile health or mHealth; [9,10]) has become increasingly common [11], and the majority of mHealth apps available in app stores [12] target weight-related behaviors such as eating behavior [13,14]. Moreover, mobile app–based interventions have been shown to be effective in improving diet and diet-related health outcomes [15] and effect sizes of these interventions are comparable to those of traditional nondigital interventions [15]. However, a large share of the general population does not yet use nutrition apps [16-18]. To promote their use and thus to put their full potential into effect, further knowledge about reasons for and against adoption and prolonged use of nutrition apps is necessary.

These reasons are likely multifaceted. For instance, research in the area of wearable health devices (eg, fitness trackers, smartwatches) has already shown a large variety of reasons for device uptake, sustained use, and abandonment related to, for example, health status, health-related goals, (de)motivation, perceived utility, measurement inaccuracy, usability, convenience/accessibility, and privacy [19,20]. However, these findings may not be easily transferrable to nutrition apps as they differ in central aspects. For instance, nutrition apps usually require manual initiation of data recording by opening the app or pressing buttons, and often also manual input of consumed foods to provide feedback, for example, by having users enter foods and estimated serving sizes using a comprehensive food database [21]. Fitness apps and wearables, by contrast, allow for automatic data collection based on various sensor technologies (eg, accelerometers or gyroscopes). Hence, the reliability of these 2 mHealth services refers to different data sources—the user versus technology—and might thus be perceived differently by users. Furthermore, the difference in active versus passive data collection might impact the motivation to use the services long term, as tracking food intake manually might require more effort, when compared with tracking physical activity automatically [17]. Finally, mHealth services for nutrition and physical activity differ in the type and temporal pattern of feedback provided. Because of automatic and continuous collection of (in)activity data, wearable fitness trackers can provide feedback continuously, even while an activity is taking place, which allows for immediate adjustments. For nutrition apps, by contrast, data (and thus feedback), are provided for distinct eating occasions. Moreover, feedback is often only available after the food was consumed or at least chosen. Thus, immediate adaptation of the behavior might not be possible, which might impact how feedback is perceived and evaluated. Therefore, barriers and facilitators for nutrition apps may differ at least in part from barriers and facilitators for the use of fitness apps and wearables. Finally, these differences might also explain why fitness apps are generally more popular than nutrition apps [17].

Two systematic reviews have identified several factors influencing engagement with mobile weight loss and weight maintenance interventions. These factors include personalization, simplicity, entertainment, usability, social support, and the presence of certain features such as self-monitoring, prompts, and feedback [22,23]. However, nutrition apps can be used for a variety of goals, including self-monitoring [24], eating healthier [25], or even gaining weight [26], which again may reflect a variety of underlying motivations including health status (*cf*. [19]) and specific needs and expectations. Furthermore, the methodological focus of the reviews was restricted. For example, Lyzwinski et al [23] included only qualitative studies. While qualitative research allows for a great richness of participants' responses, quantitative research usually comprises larger samples and may allow for formal testing of theory. Most studies included in Sharpe et al [22], by contrast, were randomized controlled trials which either tested the effect of the presence or absence of certain features (eg, social network, tailored content) on engagement indices or evaluated features of a specific intervention, which might not be generalizable to all mobile eating interventions. It was thus deemed necessary to review the literature more broadly to reflect a wider range of app-based nutrition interventions, study designs, and user perspectives. Accordingly, this present systematic review aimed to synthesize the literature on barriers to and facilitators for the uptake, continued use of, and disengagement from nutrition apps in the general population. This was done to provide a comprehensive overview of factors that hinder or promote use to be utilized as starting points for nutrition app development and optimization. It includes studies with all possible user groups (nonusers, users, ex-users, and not specified) [17] as well as both qualitative and quantitative studies on a wide range of available nutrition apps.

Methods

Protocol and Review Design

A protocol was prepared and published on the Open Science Framework [27] prior to completion of data extraction. This review reports on the generation of an overview of the evidence. The second goal (designing a questionnaire based on the results) will be presented elsewhere. Reporting is guided by the PRISMA guidelines [28] (Multimedia Appendix 1).

Data Sources and Search

We searched the following databases before May 2019: PubMed, Web of Science, PsychINFO, PSYNDEX, PsycArticles, and SPORTDiscus. For PubMed, Web PsychINFO, PSYNDEX, PsycArticles, and SPORTDiscus. A Boolean search term was used for this purpose: ((“nutrition app*” OR “diet app*” OR “weight loss app*” OR “weight control app*” OR “weight

management app*" OR "food journal" OR "health app*" OR "personal quantification" OR "quantified self" OR "personal informatics")) AND (adoption OR adherence OR abandonment OR attrition OR barriers OR motivation OR attitude OR *engagement OR "former user" OR ex-user OR "*continued use"). However, a slightly modified term was used for Web of Science due to differences in use of the asterisk (ie, we had to add "disengagement" and "discontinued use"). No restriction was placed on publication date. Moreover, we conducted backward citation search by manually screening the reference lists of included studies for additional relevant references. We also conducted a forward citation search in Google Scholar using the included studies to complement the data search.

Study Selection

Inclusion and Exclusion Criteria

Eligible studies examined factors hindering uptake or continued use (ie, barriers) or factors promoting update or continued use of nutrition apps (ie, facilitators). To be eligible for inclusion, studies had to examine adolescents (aged 13-18) or adults in the following participant groups: current users, ex-users, or nonusers of nutrition apps either for themselves or for their children. Studies with children, adolescents younger than 13, and health care providers using nutrition apps for patient support were excluded. Only empirical articles were included (ie, literature reviews, meta-analyses, and conference abstracts were excluded). All study designs including qualitative or quantitative methodologies were eligible for inclusion. Studies had to investigate general nutrition app use (ie, studies evaluating particular apps, for instance, in intervention settings, were excluded). Moreover, studies focusing more broadly on the use of health apps were included as long as they specifically stated that nutrition apps were included. Studies were only eligible if the examined apps included an assessment of diet, such as logging consumed foods (eg, studies evaluating sole weight logging apps were excluded). Further, we included only English and German articles. Finally, 2 studies known to authors from other sources were also included [29,30].

Screening

Authors LK and CA independently reviewed titles and abstracts, and, subsequently, full texts according to the inclusion and exclusion criteria. Conflicts were resolved through discussion until consensus was achieved.

Data Extraction and Collation

Authors LK and CA reviewed each full-text article independently and extracted data on facilitators and barriers to nutrition app use (ie, direct quotes from qualitative studies, questionnaire item texts from quantitative studies). Subsequently, all authors categorized facilitators and barriers manually according to principles of qualitative content analysis (ie, inductive category development) [31]. Differences in abstraction were resolved by discussion until consensus regarding category logic (ie, no overlaps of contents across categories) was achieved. The inducted categories were first defined, then compared and harmonized with the extracted quotes. The final category system including the underlying quotes was documented in an MS Excel spreadsheet and can be found in Multimedia Appendices 2 and 3. Categories may contain both barriers and facilitators as several aspects of nutrition apps and their use were not universally perceived to be positive or negative (eg, app usability can both be a facilitator and barrier depending on the individual user's perception). In a final step, the categories were grouped into higher-order clusters based on their origin or source as they might provide further insight into starting points for improvement.

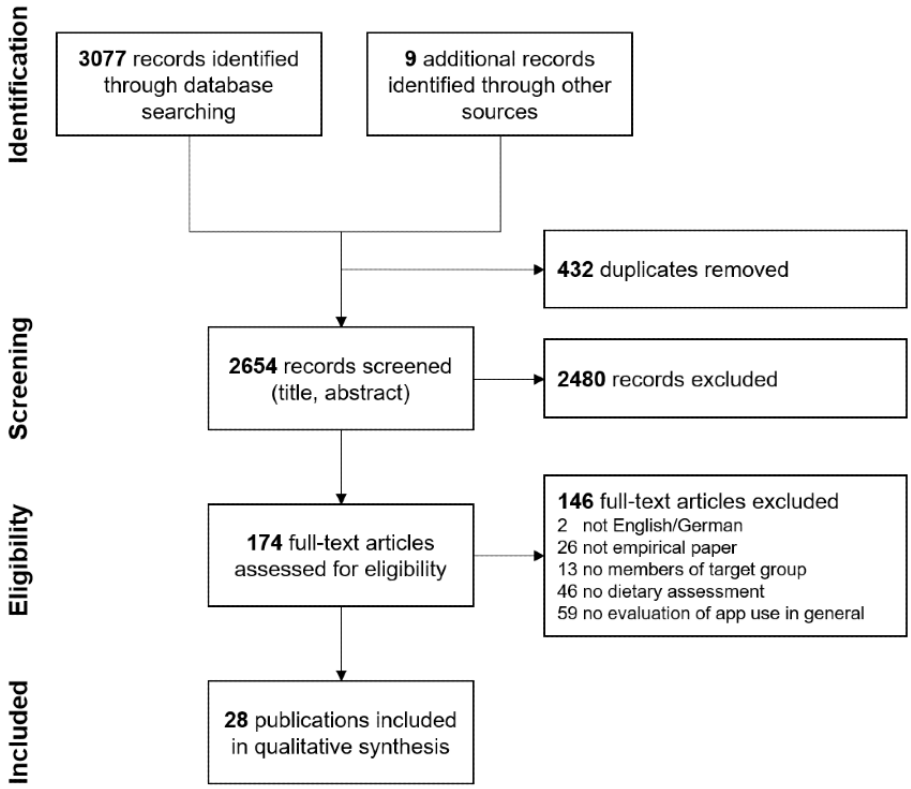
In addition, author LK reviewed each full-text article and extracted the following study characteristics into MS Excel spreadsheets: app user group (all, users, or users and ex-users), sample size, age, gender, specificities of the study sample (general population or patients), study design (qualitative, quantitative, mixed), study location, and type of app (nutrition vs health app; see Multimedia Appendix 4).

Results

Literature Search

A total of 2654 individual records were screened. After 2480 were excluded by screening titles and abstracts, 174 full texts were screened for eligibility. Subsequently, 28 publications containing 30 studies [24-26,29,30,32-54] were included (see Figure 1 for a flow diagram). Of these, 9 publications were identified through other sources: 7 were identified through forward and backward citation search [25,26,35,39,41,45,47], and 2 were known to the authors through unrelated literature searches [29,30].

Figure 1. Flow diagram of article selection.



Characteristics of the Included Studies

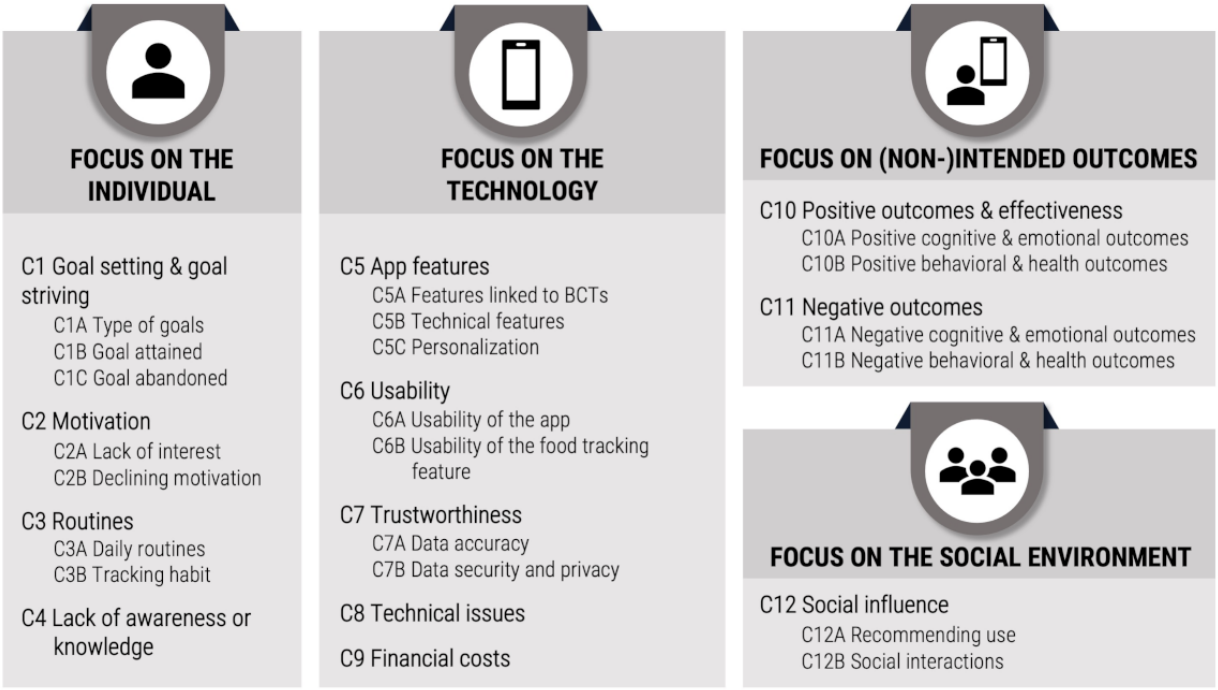
The 28 publications were all published between 2012 and 2019. Two publications contained 2 studies [49,53], while the remaining 26 publications each reported results from 1 study [24-26,29,30,32-48,50-52,54]. Almost all samples comprised adults from the general population, while 1 publication [45] focused on adolescents and 1 publication [26] focused on women with an eating disorder. Eight publications [26,33,34,36-38,44,45] focused specifically on barriers and facilitators of nutrition app use (eg, diet tracking features in MyFitnessPal, LoseIt!), while 20 publications [24,25,29,30,32,35,39-43,45-48,50-54] investigated barriers to health apps in general, but explicitly included nutrition apps such as calorie counters. Fifteen of 28 [24,25,29,30,32,38,41,43,45-47,51-54] included publications did not focus on a specific user subgroup, while 5 publications [35,36,39,48,50] focused on users and 8 [26,33,34,37,40,42,44,49] included both users and ex-users. For the majority of publications, data were collected in the United States (15/28) [24-26,29,30,32,34,35,40,41,43,45,51,52,54], followed by the UK (3/28; [37,38,47]) and Canada (2/28; [33,48]). One publication presented data from multiple countries. Ten of the 28 publications referred substantially to theoretical models for the deduction of hypotheses or discussion of results, while the remaining studies were not theory driven. See Multimedia Appendix 4 for an overview of included studies.

Barriers to and Facilitators for Nutrition App Use

A total of 328 barriers to and facilitators for nutrition app (non)use were extracted from the publications. The number of extracted barriers and facilitators varied greatly between publications (mean 11.71 [SD 8.75]; range 2-39).

While grouping barriers and facilitators into categories, a 3-level hierarchical framework emerged (see Figure 2 for a schematic overview of the framework). First, barriers and facilitators were grouped into 23 subcategories. Second, several categories were clustered (eg, lack of interest and declining motivation were both related to motivation, thus grouped together), resulting in 12 categories (C1–C12; note that C4, C8, and C9 do not contain subcategories). Third, the resulting categories were grouped into higher-order clusters that focus on either (1) the individual (C1–C4), (2) the app and the smartphone (C5–C9), (3) intended and nonintended outcomes of nutrition app use (C10 and C11), or (4) social influence (C12).

Figure 2. Framework comprising 12 categories C1-C12 and respective subcategories of barriers and facilitators identified in the individual studies. BCTs: behavior change techniques.



In the majority of study results (23/30), barriers and facilitators were not distinguished by user group (ie, no differentiation regarding barriers or facilitators for uptake, use, or long-term use). Therefore, such a differentiation is also not reflected in the presented framework. However, a differentiation regarding study samples was possible and can be inspected visually in Multimedia Appendix 2.

Barriers and facilitators within the individual were related to C1 goal setting and goal striving (ie, C1A the type of goal; C1B goal attained; C1C goal abandoned), C2 motivation (ie, C2A lack of interest; C2B declining motivation), C3 tracking routines (ie, C3A daily routines; C3B tracking habit), and C4 lack of awareness or knowledge. Barriers and facilitators related to the app and the smartphone were C5 app features (ie, C5A features linked to behavior change techniques [BCTs]; C5B technical features; C5C personalization), C6 usability of the app (C6A) and the food tracking feature (C6B), C7 trustworthiness regarding data accuracy (C7A) as well as data security and privacy (C7B), C8 technical issues, and C9 financial costs of the app. In addition, barriers and facilitators were identified that stem from using the app, which include intended and nonintended positive outcomes and effectiveness (ie, C10A positive cognitive and emotional outcomes; C10B positive behavioral and health outcomes) as well as C11 negative outcomes (ie, C11A negative cognitive and emotional outcomes; C11B negative behavioral and health outcomes). Finally, barriers and facilitators stemming from C12 social influence (ie, C12A recommendation to use; C12B social interactions) were identified.

There was a substantial variation in the number of subcategories targeted by the different publications (mean 6.79 [SD 3.87]; range 1-16 of 23 subcategories) and accordingly, the frequency of the subcategories being mentioned across publications varied substantially (mean 8.26 [SD 4.21]; range 1-19 of 28 publications; see Multimedia Appendix 2 for details).

C1 Goal Setting and Goal Striving

C1A Types of Goals

In 15 publications [24-26,29,30,32,33,35-42], the type of goal was mentioned as a facilitator for nutrition app use. A variety of goals were identified ranging from highly specific nutrition-related goals to very general health-related goals, or even to improvements in other aspects of life. The majority of goals were related to nutrition, for example, food tracking [24,26,32,33], diet improvement [25,34-36], and weight management [32] including both weight loss [26,29,35,37,38] and weight gain [26]. In addition, changing other health behaviors was mentioned [30,33,35,39-41], for example, physical activity [35] as well as adopting a new or maintaining an existing behavior [35]. Further participants named more general goals such as improving their health [42], to be more mindful or to find balance, needing assistance with medical or health-related decision making, increasing their knowledge in order to answer specific questions, finding triggers, being able to ask their physician more specific questions or to ask for a second opinion, curing or managing a condition, or executing a treatment plan [39,41]. Finally, some participants identified further goals related to other aspects of life and new life experiences, such as maximizing work performance [39]. These results highlight that nutrition apps may be used for a great variety of different goals; apps may thus need to be explicit about which goals they target to attract appropriate users and so to facilitate uptake and long-term use.

C1B Goal Attained; C1C Goal Abandoned

Two reasons for disengagement with an app which were related to goals were identified. On the one hand, app use may no longer

be necessary if the goal was reached or a desired habit was formed [29,33,43,44], which was identified in 4 publications. On the other hand, 1 publication highlighted that goals may be abandoned and the app along with it [29].

C2 Motivation

C2A Lack of Interest

In 4 publications [24,30,43,45], nonusers expressed a general lack of interest in using health apps including nutrition apps because they felt that they did not need to use one [24,30], for example, because of other available tools such as paper and pencil diaries [43], or competing interests such as preferring to use their smartphones for other apps (eg, social media) [45].

C2B Declining Motivation

Furthermore, in 7 publications [24,33,36,44,46-48], some (ex)users reported that their motivation to use apps could decline over time [24,36,44,46], for example, because of limited progression [33] or boredom [6], for instance, because the app only provided a limited range of functionalities [48].

C3 Routines

C3A Daily Routines

Four of the reviewed publications [29,32,33,47] highlighted that the fit between the user's daily routines or current living situation and app might impact uptake and adherence [33]. For instance, participants in 1 study [29] discontinued using an app because they were not able to use it in certain environments (eg, at work). Others criticized that using an app interfered with their daily activities and social life [32]. Dennison et al [47] highlighted that apps were most likely to be utilized if they were well integrated into users' typical smartphone use patterns. Thus, according to the reviewed studies, nutrition apps need to fit to the user's daily routines, for example, by not being able or not wanting to use the app for inputting data while at work.

C3B Tracking Habit

In 4 publications [40,44,47,49], tracking habits (or a lack thereof) was identified as an influencing factor. Some participants stopped using an app because they forgot to use it in daily life [44,47,49]. Similarly, Yuan et al [40] showed that apps were less likely to be abandoned when a habit of using the app was formed. Nutrition apps might therefore need to include features that facilitate establishing a tracking habit to promote their use.

C4 Lack of Awareness or Knowledge

Four publications [30,33,43,45] highlighted the importance of knowledge and skill especially for nutrition app uptake. Some of the nonusers surveyed in the reviewed studies reported not to be aware that health and nutrition apps existed [30,43]. Others may be aware of this type of apps, but are unsure which one to use, lack awareness of specific functionalities and capabilities, or do not know how to use them properly [33,43,45].

C5 App Features

C5A Features Linked to Behavior Change Techniques

In a total of 14 of the included publications [29,33,36-39,42,43,45-50], the inclusion of features that can be linked to the BCT Taxonomy version 1 [55] was evaluated as a positive factor. Participants explicitly expressed their satisfaction with comprehensive food databases for self-monitoring (BCT category 2) of food intake [37,48] and criticized databases that were too limited [49]. Regarding feedback, the opportunity to view a history of tracked data without visiting a medical doctor or having the opportunity to send data to health professionals remotely [50] was appreciated. Similarly, participants in the study by Aljuraiban [46] were more likely to discontinue use if monitoring by a specialist was not offered. By contrast, there were disagreements on how messages should be designed [33]: some participants stated that they did not like to count calories [38] or to restrain themselves based on feedback from the app [47]. Regarding information presentation, participants suggested using various media formats (eg, video, audio) [45] and visualizations [42]. They were keen on the fact that a high level of detail was preserved in the feedback [47] and preferred to access information on the go [37]. Participants generally valued the provision of nutrition knowledge (BCT category 4) that they did not already have or might not be able to access otherwise [36,38,43]. Finally, rewards (BCT category 10) were appreciated [37,39,42,43,45]; however, gamification elements were seen favorably by some [36,43] and perceived as demotivating by others [36].

C5B Technical Features

In 9 studies [33,36,37,42,43,47,48,50,51], participants mentioned the inclusion of further technical features. For instance, participants highlighted the need for integration with other apps, for example, to synchronize calorie consumption and expenditure [33,42]. Poor integration with other apps might lead to disappointment with an app and to subsequent disengagement [33]. Integration of location tracking was criticized as unnecessary in Dennison et al [47]. However, based on the included quotes, one can only speculate whether these features will actually lead to abandonment of an app or simply be ignored. Finally, receiving messages, prompts, and reminders to use the app were not universally appreciated. While some participants stated that they were helpful [36,37,43,47], others reported to be annoyed by too frequent notifications [33,47,48,51], for example, to update the app regularly to ensure its functionality [50].

C5C Personalization

Ability to personalize apps was mentioned in 11 publications [29,33,36-38,43,45-48,50]. Features should be customizable and tailored to individual needs and goals [29,36-38,43,46,50]. For instance, participants valued the opportunity to set customized reminders, to have choices in message content and tone [47], and to receive personalized information and coaching through an app [45]. Some participants stated liking apps that provided a prespecified list of goals to choose from [29,33], while others stated that they would prefer to set individual goals as they were unsatisfied with the ones provided [33,36]. In

addition, tailoring to ethnic and age-specific (eg, adolescence) preferences was valued [45,48].

Moreover, it might be valuable to tailor the features implemented in an app to the users' needs. In several studies, the absence of desired or helpful features [33,42,48] was mentioned, which may lead to abandonment of the app [29]. By allowing users to customize the app, personal autonomy, freedom of choice, and a feeling of congruency are preserved, which might prolong use [36]. Thus, as needs and expectations regarding the design and content of certain app features may vary between users, customization or variation of both features and app content in general may thus be helpful to prevent abandonment.

C6 Usability

C6A Usability of the App

In 19 publications [24,25,30,33-39,43,45-50,52,53], usability of the app was seen as an important precursor of nutrition app uptake and continued use. Participants criticized that apps were too confusing, complex, and stressful, both when setting them up (eg, because of lengthy instructions) [35,47] and when using them [24,33,35,38,43,45-49]. Accordingly, easy and simple tracking procedures were valued by users [39] and perceived ease of set up and use increased the likelihood for adoption and continued use [25,30,33,34,36,37,50]. In a similar vein, participants reported disliking apps that were time-consuming to use [36,47,52], which might lead to increased levels of stress [53]. Lack of time might thus be considered an important barrier to using nutrition apps [43]. Finally, interface design aspects were also mentioned in a few studies. Specifically, studies highlighted that an attractive design may increase the likelihood for an app to be used [37,45].

C6B Usability of the Food Tracking Feature

Food tracking features are an integral part of nutrition apps. The importance of their usability was highlighted in 7 of the reviewed studies [24,33,39,42,44,48,49]. Especially, the ease of use of food databases was seen as a critical component. Many users of nutrition apps reported to have experienced difficulties when entering data [44,48], for example, regarding the correct identification of foods because of too many options [33], finding the correct foods because they were missing from the database [49], or entering the correct foods and portion sizes [33]. Providing detailed entries and entering homemade food or meals consumed at a restaurant were seen to be especially challenging [49]. Participants might even cease to enter their meals, for instance, when consuming a variety of foods over a longer time span [49]. By contrast, Lieffers et al [33] noted that participants preferred larger over smaller food databases, as these saved time and were convenient.

Furthermore, the time needed for food journaling was seen as crucial by participants [33,42,49]. Entering foods being too time-consuming was named as an important reason for not using or ceasing to use nutrition apps [24,44]. Participants would therefore prefer automated tracking functions [39] or food scanners [37]. However, Lieffers et al [33] noted that several participants preferred using an app-based food database over other nondigital tracking methods because apps were perceived as more convenient and less time-consuming.

C7 Trustworthiness

C7A Data Accuracy

A (lack of) concern regarding data accuracy was identified in 7 publications [33,36,37,42,47-49]. While some participants saw nutrition apps as a trusted ally that supported them in achieving their goals [36], other participants were concerned regarding the accuracy and trustworthiness of information presented in the app [42]. Some participants expressed concern regarding human error when tracking food intake, as tracking tools might allow deliberately adjusting entries, which may lead to inaccurate records [47]. Other participants stated being concerned about tracking errors within the database or the app itself [47,49], for example, missing or duplicate food entries or incorrect caloric information [33], or about being unsure whether the provided information, for example, in discussion boards, was accurate and could be trusted [37,47]. Moreover, participants criticized that apps may be misleading regarding the predicted accuracy of provided information [48].

C7B Data Security and Privacy

A number of data security concerns were mentioned by participants in 10 of the reviewed publications [7,24,29,35,36,42,43,47,50,51]. Some participants were unsure what the data might be used for without their awareness [47]. They expressed worry that their potentially sensitive data would be made available to third parties such as health insurers [50] or companies, for example, to tailor advertisements to them [47], although they would not consent to the data being shared as these are private matters and could be exploited [43]. In particular, location detection was seen critical because it might be a risk to personal safety, for example, when GPS data were accessed by burglars [42,47]. Zhou et al [51] specifically focused on data security concerns and barriers and facilitators to the use of mHealth apps including nutrition apps. Concerns were raised because of storing unencrypted personal data on users' smartphones, sending data to remote servers without permission, and a general lack of privacy statements in many apps. Accordingly, participants were unsure whether they could trust the apps and their developers [24,29], and apps provided by health experts such as medical doctors were seen as more trustworthy and persuasive [47]. In one study, participants expressed that they would trust apps by organizations more than apps by commercial companies, and that they would trust apps that were branded or labeled [43]. Indicators for a credible source may thus be important for nutrition app uptake. Other participants, however, stated to be unconcerned because they thought their data would not be of interest to third parties [47].

Especially in apps with in-app community features or connections to social media, anonymity was important to app users [36]. However, there were different opinions as to whether participants approved of sharing data with a community: Some current users were favorable toward sharing their data online [42], whereas for others, sharing data without one's awareness was a frequently named reason for abandoning an app [24]. Sharing might be disliked especially if participants feel that

they cannot control which and how much information is shared [47]. Besides ensuring that data are stored safely and privacy is respected, it may thus also be important to make data protection efforts transparent to potential users.

C8 Technical Issues

An important prerequisite for being able to use nutrition apps is compatibility of the app with one's smartphone [29]. However, even if the smartphone fulfills an app's basic system requirements, a number of technical issues might decrease the likelihood of app adoption and continued use. Accordingly, technical issues were identified as a potential barrier in 8 publications [24,29,33,42,43,47,48,51]. Apps may slow down the smartphone [51] and thus may impair the use of other apps or phone functions [47] or lead to crashes [33]. Furthermore, excessive battery drain [29,43,47], use of memory or storage [43,47], and use of mobile data [24] were criticized in several studies. Moreover, technical issues within the app such as app dysfunctions [48] and inconvenient data transfer [42] may lead to frustration and subsequent disengagement. Finally, users might disengage from using an app because of concerns related to radiation from their smartphones [43]. App developers therefore might need to find a balance between using advanced but resource-intensive features and ensuring compatibility with many different smartphones that vary in age and technical features.

C9 Financial Costs

Financial costs of nutrition apps were mentioned in 8 publications [24,30,33,35,40,43,46,51]. While few users may be more motivated to use the app because they paid for it [33], the price and costs of in-app purchases mainly hindered app use [24,30,43], especially because there are free apps available [35,51]. Some participants, however, said that they might be willing to pay some money for the app if it was good and provided a good value for money [35,40]. Furthermore, some participants criticized hidden costs (eg, for enabling additional app features) [46], which may lead to disengagement [24]. Thus, financial costs might need to be low and made transparent to promote nutrition app use.

C10 Positive Outcomes and Effectiveness

C10A Positive Cognitive and Emotional Outcomes

In 13 publications [25,29,32,34-38,40,42,43,50,53], participants reported a number of positive outcomes of nutrition app use that were related to cognitions and emotions. Regarding cognitive outcomes, using an app may increase users' awareness and motivation for healthy eating [25,34,35,37,43,50,53] and induce positive feelings, including feeling energized and healthy [29,32]. Accordingly, nutrition apps are perceived to be informative and to promote nutrition knowledge [38,53]. Tracking may also provide participants with a sense of accountability and the ability to track progress over time [29] and improve their self-efficacy [36]. Regarding emotional outcomes, participants reported that using an app made them feel good about themselves [53] and their bodies [32]. Furthermore, apps may provide encouragement and support [35,38], and using them can be fun and enjoyable [34,40] which may positively influence engagement [42].

C10B Positive Behavioral and Health Outcomes

Positive impact of nutrition app use on behavior and health [25,34], potentially due to improved self-management skills [50], was reported in 6 publications [25,30,34,47,50,54]. Accordingly, nutrition apps are seen as potentially useful [34,54], as merely viewing behavioral data may facilitate behavior change [47]. However, not all surveyed participants agreed with this notion and perceived nutrition apps to not be effective in changing health and related behaviors [30,47].

C11 Negative Outcomes

C11A Negative Cognitive and Emotional Outcomes

A number of potential adverse cognitive and emotional consequences of app use were reported in 8 studies [32,33,37,38,42,47-49]. Participants reported obsession with food or calorie counting [32,33,38,49] and being overly engaged with one's states and behaviors [42]. Moreover, negative app-generated information including feedback based on tracking of food intake and messages sent by the app might evoke negative emotional reactions including disappointment, guilt, and anxiety [32,33,37,48,49], especially if users fall short of reaching a predetermined goal [47]. Finally, apps may also make users feel neurotic about their body image [32].

C11B Negative Behavioral and Health Outcomes

Using nutrition apps might also have a negative impact on behavior and subsequently health, which was reported in 7 studies [26,33,43,44,48,49,52]. Some participants expressed worry that eating foods that are unhealthy, but easy to log (eg, ready meals) would be rewarded, which may encourage their consumption [44,49]. Other participants expressed concern that feedback on calorie consumption might backfire, for example, when caloric intake and expenditure are tracked in combination and burned calories may be seen as a permission to eat more [33]. Finally, some participants stated being concerned that apps promoting extreme calorie restriction might even lead to potential harm [48], including inducing or exacerbating an eating disorder [26]. Nutrition app use might therefore be promoted if expected or experienced negative consequences are attenuated.

C12 Social Influence

C12A Recommending Use

Social influence might promote app uptake, as was highlighted in 9 publications [24,29,30,35,43,45,48,51,54]. For instance, participants may learn about apps from family members or friends [29,30,35,43], from a health or fitness professional, or from their employer [24,29]. Still others stated to have chosen the app based on recommendations and positive reviews in app stores, social media, or TV [29,30,54].

C12B Social Interactions

Similarly, 11 publications [33,42-45,47-49,51,53,54] indicated that social interactions in the app or related to the app might influence nutrition app use. For instance, participants valued competitions [43,48] and support functions [37]. Furthermore, Wang et al [53] reported that sharing data over the internet might sustain motivation in users. When their friends stopped using the app, participants reported that it was more likely that they

would stop using the app as well [49]. However, some participants also reported that learning about others' success or competing against other app users might be demotivating [33,43,49].

By contrast, perceived undesirability and stigmatization of using apps might hinder uptake and continued use [45,47,49] because it may be embarrassing [47]. Some participants stated that they did not feel comfortable using an app in front of others [33,44] or did not even want other people to know that they were using an app because it might imply that they have a certain disease [51]. Thus, social influence might both promote and hinder nutrition app uptake and use, depending on whether attitudes toward nutrition apps in the social environment are positive or negative.

Discussion

Principal Findings

Nutrition apps are less popular than fitness apps [17], although they might have comparably beneficial effects on health [15,56]; for instance, on body weight reduction [15]. Hence, to enable large-scale health effects such as tackling the obesity epidemic [57], acceptance, wide-spread adoption, and long-term use of nutrition apps need to be enhanced. To this end, it is necessary to better understand differences and dynamics in use. This systematic review provides a hierarchical framework of barriers and facilitators for nutrition app use. The framework highlights that besides technological reasons, characteristics of the (potential) user, the interplay between user and technology, and the social environment impact whether a nutrition app is used. For instance, it underlines the importance of tailoring the app content to the user's goals, expectations, and needs. As Villinger et al [15] pointed out, nutrition apps mainly employ 4 categories of BCTs [55,58] that primarily address constructs of deliberate behavioral control, such as goal setting, self-monitoring, and feedback (see also [59] for an analysis of commercial apps). However, as some (prospective) nutrition app users may decide what or how much to eat based on intuition rather than based on deliberation, the number of nutrition app users might be increased, for instance, by developing apps that address a preference for intuition [17].

The broad range of barriers and facilitators identified in this review may be due the great variety of samples and study designs included that exceeds previous reviews on the topic. For instance, 9 of the 11 studies included in Sharpe et al [22] were evaluations of randomized controlled trials or ongoing weight management programs. Researchers might thus have been especially interested in usability evaluations of specific program features that support long-term engagement with the tested intervention. Consequently, their synthesis puts an even stronger emphasis on barriers and facilitators related to technology. When including surveys of general population samples that are not restricted to participation in an intervention study, more barriers and facilitators may be identified that might not play a role for study participants. For example, lack of awareness or knowledge might not emerge as a barrier when evaluating a digital weight management program because participants often attend a training session before the start of the study (eg, [60]). Similarly, technical issues might not be as frequent, as participants might be preselected based on the type of smartphone they use (eg, [61]), or they might receive a smartphone from the study team to ensure compatibility (eg, [62], Study 3). Besides, data security concerns might not be of great importance when taking part in a study at a university, as researchers might be seen as more trustworthy [47]. Finally, financial costs might not play a role because using an app as part of a study is usually free. Thus, by explicitly including studies independent of the use of specific nutrition apps, this review was able to generate a comprehensive list of barriers and facilitators that play a role when deciding whether to use (or continue to use) a nutrition app in daily life.

Previous reviews also often focused on specific user groups such as current users of nutrition apps [22] or remote tracking technology [19] or on ex-users of wearables [20], and thus lack the perspective of potential users who did not yet think about using mHealth technology for health promotion or decided against its use (for a discussion, see also [17]). This review, by contrast, included literature on nutrition apps in general and included current and ex-users of nutrition apps as well as nonusers of nutrition apps and thus generated an extensive list of barriers and facilitators. Consequently, several subcategories were predominantly or even exclusively identified in publications which included ex-users or nonusers in addition to current nutrition app users. For instance, the category C4 (lack of awareness or knowledge) was not mentioned in publications that exclusively focused on current users, and the subcategory C2A (lack of interest) was only identified in publications which also included nonusers. However, one could argue that these barriers are especially important to address in order to facilitate contemplation of the use of nutrition apps (eg, through medical prescription which is now possible in Germany [63]), which is an important first step in the nutrition app adoption process [17,64]. Similarly, the category C3 (routines) with its subcategories C3A (daily routines) and C3B (tracking habit) was not identified in studies which focus exclusively on current users, presumably because a significant number of current users have already established a tracking habit and are using an app that fits their daily routines. Although certainly not all current users of mHealth technology already have established a tracking habit, those that do are likely to use the technology for periods of a year or more [20,65]. Thus, forming a habit of using a nutrition app might be an important prerequisite for prolonged nutrition app use (see [66] for a discussion of the importance of habits for behavior change). To be able to take habit formation into account when developing nutrition apps, further research is needed to identify app features (eg, reminders, identification of event-based cues [67]) that are most beneficial for establishing a tracking habit.

Despite the differences in target populations and technologies, there is substantial overlap between barriers and facilitators to different mHealth technologies identified in this review and previous reviews [19,20,22,23]. For instance, reviews on both nutrition apps and wearables identified facilitators and barriers related to data security and privacy, app features, or technical issues. Thus, many of the identified barriers and facilitators could be generalized across mHealth technologies, and the

presented framework may provide insights for designing mHealth technologies more broadly.

Implications for the Development of Nutrition Apps

The framework developed in this review summarizes barriers and facilitators for nutrition app use based on empirical investigations and can therefore guide development of theory and measurement instruments, for example, questionnaires to assess interindividual factors related to app uptake. Moreover, the identified barriers and facilitators, specifically on the level of technology (C5–C9) as well as outcomes of user–technology interaction (C10–C11), can inform the improvement of nutrition apps. To this end, 8 design guidelines derived from the presented results are listed in Table 1. First, usability was the most frequently identified barrier in this review and was also identified in previous reviews [19,20,22,23]. Importantly, usability issues were identified both concerning the app in general (design guideline 1; eg, navigation through the app) and more specifically concerning the food tracking feature and the underlying food database (design guideline 2). The latter constitute core features of many nutrition apps [68,69]. Previous research has shown that usability issues related to tracking of food intake might impact willingness to record eating events [70]. Therefore, deficient usability of the food database might also indirectly impact other categories of barriers and facilitators identified in this review such as accuracy and trustworthiness if fewer meals are recorded. Subsequently, this may also impact features that rely on accurate data, such as feedback. Usability of nutrition apps and especially food tracking features should therefore be a major concern for app developers. User burden can, for example, be reduced by using simpler input mechanisms, such as by indicating portion sizes using common household items [71], or by photo-based food recording [68,72].

Table 1. Design guidelines based on review results.

Design guideline	Related references
DG1[a]: Enhance app usability through quick set up, avoidance of lengthy instructions, and high ease of use.	[24,33,35,38,43,45-49]
DG2: Enhance usability of the food tracking feature through simple input mechanisms and enabling quick and reliable identification of the correct foods (eg, by including automated tracking functions or barcode scanners, indicating portion sizes, photo-based food recording).	[24,33,37,39,42,44,48,49,68,71,72]
DG3: Include effective behavior change techniques, for example, goal setting, self-monitoring and providing feedback, rewards, and shaping knowledge.	[22,23,36,37,48,73-75]
DG4: Allow for personalization of the app to fit individual needs and goals (eg, give choice to enable or disable gamification elements or reminders, offer customizable reminders, and adaptability to personal variables such as ethnic preferences or age group).	[19,22,23,29,33,36-38,43,45-48,50,59,76]
DG5: Anticipate possible outcomes of nutrition app use to promote positive outcomes (eg, increase of awareness and motivation for healthy eating, improvement of self-management skills, and nutrition knowledge) and avoid negative outcomes (eg, obsessive calorie counting, feelings of guilt, disappointment, or anxiety).	[23,25,29,32,34-38,40,42,43,47,50,53,54,77-79]
DG6: Advance trust in data accuracy by restricting opportunities for human error (eg, when tracking food intake) and enhancing data transparency (eg, specify source of nutritional values).	[33,37,42,47-49]
DG7: Enhance data authority by providing transparency regarding data sharing with companion and third-party apps and giving the choice to prohibit data transfer.	[24,29,42,43,47,48,51]
DG8: Be economical regarding use of smartphone resources (eg, avoid excessive memory and mobile data usage and battery drain).	[24,29,33,43,47,51]

[a]DG: design guideline.

The presence of certain app features including self-monitoring and feedback features can be seen as a facilitator for continued nutrition app use (see also [22]). Specifically, in this review, several app features could be identified that can be subsumed under BCT categories included in the BCT Taxonomy [55], such as self-monitoring (BCTs 2.3 and 2.4), feedback (BCTs 2.2 and 2.7), or goal setting (BCTs 1.1 and 1.3). As Lyzwinski et al [23] pointed out, the inclusion of BCTs is often valued by nutrition app users. It could thus be concluded that the inclusion of certain BCTs (design guideline 3) might not only increase effectiveness of interventions [73,74] but also engagement [75]. Moreover, both this review and previous reviews highlighted that users appreciated opportunities for personalization of the app [19,22,23], which is also related to effectiveness in the literature [59,76]. It can therefore be recommended to include features such as feedback, goal setting, and prompting, and to allow for customization, for example, by allowing users to set customized reminders, to increase engagement (design guideline 4). Many BCTs (eg, 4.3 reattribution, 4.4 behavioral experiments, 5.2 salience of consequences), however, were not mentioned by the study participants. Future research therefore needs to investigate their effects on nutrition app uptake and prolonged use.

Moreover, anticipated or experienced positive and negative outcomes, including the (lack of) effectivity, were among the most frequently identified reasons for nutrition app (non)use. Interestingly, positive and negative outcomes of use were only rarely addressed in previous reviews. For example, potential negative consequences of using nutrition apps such as feelings of guilt were only explicitly addressed in Lyzwinski et al [23]. From a psychological point of view, however, anticipated or

perceived positive and negative consequences for one's health are important precursors of engaging in a behavior (eg, [77,78]). Proximal outcomes, which include cognitive and emotional consequences of use, might be especially important for continuously performing a behavior [77]. Accordingly, outcome expectancies are central components of models of (health) behavior including Social Cognitive Theory [80] and the Health Action Process Approach [81]. It could therefore be recommended to anticipate potential negative outcomes of using an app to prevent them. At the same time, potential positive consequences of nutrition app use, including positive emotional consequences (eg, increased well-being [79]) should be emphasized more to promote use (design guideline 5).

Another critical factor influencing the acceptance of nutrition apps is data accuracy. In the area of nutrition apps, this factor is closely related to the data quality of the underlying data base, which opens up room for human error by allowing users to add their own entries. Consequently, incorrect nutritional values and uncertainty regarding their source might follow [33,37,42,47-49]. If human error cannot be completely avoided, opportunities should be created to increase transparency regarding data sources and thus trust in the system (design guideline 6). Furthermore, barriers in the area of privacy protection were identified [24,29]. In order to strengthen the data sovereignty of users, concerns regarding the transfer of data to third parties should be addressed by making them transparent and optional (design guideline 7). Finally, aspects of sustainability and energy efficiency play a role in the acceptability of nutrition apps [29,43,47]. Manufacturers should thus take care to design their apps sparingly in terms of data storage and energy use (design guideline 8).

Previous reviews have already highlighted a lack of theory when developing the content of app-based interventions (eg, [56,82]). However, when aiming to understand factors related to engagement with nutrition apps, less than half of the publications included in this review used theories such as the technology acceptance model [83] or the theory of planned behavior [84] to design the study or interpret its results. As the links between the identified barriers and facilitators and existing frameworks and models of health behavior highlight, psychological theory may be highly beneficial to gain a better understanding of engagement with health apps and the design of more engaging apps [85]. It is therefore important to use theory in future studies about health app uptake and prolonged use as well as to use theory to inform app development.

Limitations and Avenues for Future Research

Some concerns regarding the findings of the review arise from the included studies. Several of the included studies investigated reasons for health app (non)use more broadly. Although only studies were included that specifically addressed nutrition apps as a category of health apps, it could often not be determined whether an issue was raised in relation to nutrition apps or other categories of health apps. Thus, some barriers and facilitators might have been included in this review that did not refer to nutrition apps. Furthermore, most of the included studies did not provide information about anthropometrics or socioeconomic position of the participants, which makes it difficult to appraise generalizability of the findings. Although most studies reported the gender of the participants, females were overrepresented in many of the included studies, and 2 even focused exclusively on female participants [26,36]. While previous research suggests that nutrition app users are more likely to be female [17], including more male participants in future research would be desirable to address potential gender-specific needs (eg, aiming to lose weight vs aiming to gain muscle mass [86]), which might explain the lower adoption rates in males. Finally, selection bias, for example, due to publication bias, cannot be ruled out in both quantitative and qualitative studies [87]. It thus cannot be ruled out that relevant (unpublished) work may have been missed.

It is important to note that neither this review nor previous reviews on barriers and facilitators for nutrition app use can provide insights into the relative importance of the barriers and facilitators for the decision (not) to use a nutrition app. While the number of studies in which a reason was mentioned could be used as an indicator, it might also reflect research questions or questions and items used in the individual studies. Moreover, the grouping of barriers and facilitators into categories and clusters is somewhat arbitrary. Some categories in the framework might not be fully mutually exclusive, as for instance affinity for technology might also influence the perception of usability [88]. Notably, although Simblett et al [19] only identified 5 categories of barriers and facilitators for the use of remote tracking technologies, many of the underlying barriers and facilitators could also be mapped onto the framework presented in this review. Further research is thus needed to gain insight into interrelations of the identified barriers and facilitators and their grouping based on empirical data as well as to determine their relative importance (see, eg, [20] for wearables).

Furthermore, barriers and facilitators might differ between user groups, as has been highlighted in previous research on stage theories of behavior (eg, [64,89]). Differences between user groups could not be disentangled in this review as most studies reported barriers and facilitators for multiple user groups without indicating by which user group they were mentioned. One exception is the survey conducted by Murnane et al [29], which showed that current users experienced positive consequences of health app use such as feeling more healthy and energized, while apps were abandoned because they did not function properly or lacked desired features. Furthermore, the importance of barriers and facilitators might change while a nutrition app is used (see [90] for a discussion). As Baretta et al [91] showed in the context of apps for physical activity promotion, some features such as peer and coaching support might be more important for initial uptake, while, for example, proactive motivational features are more important for promoting continued use. Similarly, Sharpe et al [22] highlighted that usability might be more important for sustained engagement with nutrition apps. Future research should therefore explicitly compare different user groups and stages to provide valuable insights into how to promote uptake and continued use of nutrition apps by specifically targeting relevant barriers and facilitators.

Moreover, at least some of the barriers and facilitators identified in this review might not be specific to the use of nutrition apps, but related to changing eating behavior independent of the mode of delivery. For instance, from some of the included publications, it did not become clear whether a lack of interest in nutrition apps referred to the app itself or changing the behavior. More research is therefore needed to disentangle these effects.

Conclusions

Through this systematic review, the literature on barriers to and facilitators for the uptake and continued use of nutrition apps was synthesized to provide a comprehensive overview of factors that hinder or promote use. A total of 328 barriers and facilitators were extracted from 28 publications and systematized in a framework with 23 subcategories clustered in 12 categories. Four higher-order clusters were formed that subsume barriers and facilitators related to technology, the individual, their interactions, and the social environment. Eight design guidelines were derived from the framework which app developers may implement to increase and prolong nutrition app use: enhance app usability, enhance food database usability, include effective BCT features, allow for personalization, anticipate positive and negative outcomes, advance trust in data accuracy, enhance data authority, and conserve smartphone resources. These design guidelines might be fruitful to support the aim of the European Union [92,93] to make web-based health promotion, including nutrition apps, more effective, user-friendly, and widely acceptable, and might ultimately contribute to achieving large-scale health effects.

Acknowledgments

This research was part of the SMARTACT project which was funded by the Federal Ministry of Education and Research (Grant No. 01EL1420A, granted to BR and Harald Schupp, Department of Psychology, University of Konstanz). Furthermore, it was supported by a peer-mentoring grant from the Subdivision Health Psychology in the German Psychological Association awarded to LMK and CA. The research was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 416228727—SFB 1410. We thank Mourad Zoubir for his valuable support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist for systematic reviews.
[DOC File , 64 KB - mhealth_v9i6e20037_app1.doc]

Multimedia Appendix 2
Identified categories per publication.
[DOCX File , 29 KB - mhealth_v9i6e20037_app2.docx]

Multimedia Appendix 3
Quotes extracted from publications including assigned category.
[XLSX File (Microsoft Excel File), 28 KB - mhealth_v9i6e20037_app3.xlsx]

Multimedia Appendix 4
Detailed overview of the included studies.
[DOCX File , 16 KB - mhealth_v9i6e20037_app4.docx]

References

1. World Health Organization. Controllng the global obesity epidemic. URL: https://www.who.int/nutrition/topics/obesity/en/ [accessed 2021-06-13]
2. Mitchell NS, Catenacci VA, Wyatt HR, Hill JO. Obesity: Overview of an Epidemic. Psychiatric Clinics of North America 2011 Dec;34(4):717-732. [doi: 10.1016/j.psc.2011.08.005]
3. Borrell LN, Samuel L. Body Mass Index Categories and Mortality Risk in US Adults: The Effect of Overweight and Obesity on Advancing Death. Am J Public Health 2014 Mar;104(3):512-519. [doi: 10.2105/ajph.2013.301597]
4. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The Disease Burden Associated With Overweight and Obesity. JAMA 1999 Oct 27;282(16):1523-1529. [doi: 10.1001/jama.282.16.1523]
5. Lim S, Vos T, Flaxman A, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 2012;380(9859):2224-2260. [doi: 10.3410/f.719894684.793533485]

6. Withrow D, Alter D. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obesity reviews 2011;12(2):131-141. [doi: 10.1111/j.1467-789x.2009.00712.x]
7. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margini C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 2014;384(9960):2107-2108. [doi: 10.1016/s0140-6736(14)62367-9]
8. Statista. Number of smartphone users worldwide from 2016 to 2021 (in billions). de.statista.com. 2019. URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2019-08-28]
9. Abroms L, Padmanabhan P, Evans W. Mobile phones for health communication to promote behavior change. eHealth applications: Promising strategies for behavior change 2012:147-166. [doi: 10.4324/9780203149096-17]
10. Nacinovich M. Defining mHealth. Journal of Communication in Healthcare 2013 Jul 18;4(1):1-3. [doi: 10.1179/175380611x12950033990296]
11. Müller AM, Maher CA, Vandelanotte C, Hingle M, Middelweerd A, Lopez ML, et al. Physical Activity, Sedentary Behavior, and Diet-Related eHealth and mHealth Research: Bibliometric Analysis. J Med Internet Res 2018 Apr 18;20(4):e122. [doi: 10.2196/jmir.8954]
12. IQVIA Institute. The Growing Value of Digital Health. Durham, NC: IQVIA Institute; 2017.
13. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]
14. Aitken M, Gauntlett C. Patient Apps for Improved Healthcare: From Movelty to Mainstream. Parsnippany, NJ: IMS Institute for Healthcare Informatics; 2013.
15. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: A systematic review and meta-analysis. Obes Rev 2019 Oct;20(10):1465-1484. [doi: 10.1111/obr.12903] [Medline: 31353783]
16. Conway N, Campbell I, Forbes P, Cunningham S, Wake D. mHealth applications for diabetes: User preference and implications for app development. Health Informatics J 2016 Dec;22(4):1111-1120. [doi: 10.1177/1460458215616265] [Medline: 26635324]
17. König LM, Sproesser G, Schupp HT, Renner B. Describing the Process of Adopting Nutrition and Fitness Apps: Behavior Stage Model Approach. JMIR Mhealth Uhealth 2018 Mar 13;6(3):e55 [FREE Full text] [doi: 10.2196/mhealth.8261] [Medline: 29535078]
18. Ernsting C, Dombrowski SU, Oedekoven M, O'Sullivan JL, Kanzler M, Kuhlmey 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] [Medline: 28381394]
19. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health: Systematic Review and Content Analysis of Findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
20. Attig C, Franke T. Abandonment of personal quantification: A review and empirical study investigating reasons for wearable activity tracking attrition. Computers in Human Behavior 2020 Jan;102:223-237. [doi: 10.1016/j.chb.2019.08.025]
21. Fuchs K, Haldimann M, Vuckovac D, Ilic A. Automation of data collection techniques for recording food intake: a review of publicly available and well-adopted diet apps. : IEEE; 2018 Presented at: 2018 International Conference on Information and Communication Technology Convergence (ICTC); 17-19 Oct. 2018; Jeju, Korea (South). [doi: 10.1109/ictc.2018.8539468]
22. Sharpe EE, Karasouli E, Meyer C. Examining Factors of Engagement With Digital Interventions for Weight Management: Rapid Review. JMIR Res Protoc 2017 Oct 23;6(10):e205 [FREE Full text] [doi: 10.2196/resprot.6059] [Medline: 29061557]
23. Lyzwinski LN, Caffery LJ, Bambling M, Edirippulige S. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. J Telemed Telecare 2017:290-302. [doi: 10.1177/1357633X17692722] [Medline: 28181859]
24. 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] [Medline: 26537656]
25. Jones M, Maddox J, Benavides-Espinoza C, Finnicum P. Use of Fitness and Nutrition Applications by College Students. Missouri Journal of Health, Physical Education, Recreation & Dance 2017;27:63-71.
26. Eikey E, Reddy M. It's Definitely Been a Journey: A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. : ACm; 2017 Presented at: 2017 CHI conference on human factors in computing systems; 6-11 May 2017; Denver, CO p. 642-654. [doi: 10.1145/3025453.3025591]
27. König LM, Attig C. Work package 1: Conducting a scoping review on barriers to and facilitators for using nutrition apps. 2019 Jun 12. URL: https://osf.io/64d2h/ [accessed 2021-06-06]
28. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009 Jul 21;6(7):e1000097. [doi: 10.1371/journal.pmed.1000097]
29. Murnane E, Huffaker D, Kossinets G. Mobile health apps: adoption, adherence, and abandonment. : ACM; 2015 Presented at: UbiComp/ISWC'15; 7-11 Seotember 2015; Osaka, Japan. [doi: 10.1145/2800835.2800943]
30. Haithcox-Dennis M, Brinkley J, Richman A, DeWeese A, Byrd IJ. Mhealth on campus: assessing undergraduates attitudes and utilization of mobile health applications. Global Journal of Health Education and Promotion 2012;15(1):134-144.

31. Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. Klagenfurt: SSOAR Open Access Repository; 2014.
32. Sarcona A, Kovacs L, Wright J, Williams C. Differences in Eating Behavior, Physical Activity, and Health-related Lifestyle Choices between Users and Nonusers of Mobile Health Apps. American Journal of Health Education 2017 Jul 11;48(5):298-305. [doi: 10.1080/19325037.2017.1335630]
33. Lieffers JRL, Arocha JF, Grindrod K, Hanning RM. Experiences and Perceptions of Adults Accessing Publicly Available Nutrition Behavior-Change Mobile Apps for Weight Management. J Acad Nutr Diet 2018 Feb;118(2):229-239.e3. [doi: 10.1016/j.jand.2017.04.015] [Medline: 28625662]
34. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling Your "App"etite: How Diet and Nutrition-Related Mobile Apps Lead to Behavior Change. JMIR Mhealth Uhealth 2017 Jul 10;5(7):e95 [FREE Full text] [doi: 10.2196/mhealth.7410] [Medline: 28694241]
35. Gowin M, Cheney M, Gwin S, Franklin Wann T. Health and Fitness App Use in College Students: A Qualitative Study. American Journal of Health Education 2015 Jul 06;46(4):223-230. [doi: 10.1080/19325037.2015.1044140]
36. Flaherty SJ, McCarthy MB, Collins AM, McAuliffe FM. A different perspective on consumer engagement: exploring the experience of using health apps to support healthier food purchasing. Journal of Marketing Management 2019 Feb 18;35(3-4):310-337. [doi: 10.1080/0267257x.2019.1576756]
37. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
38. Solbrig L, Jones R, Kavanagh D, May J, Parkin T, Andrade J. People trying to lose weight dislike calorie counting apps and want motivational support to help them achieve their goals. Internet Interv 2017 Mar;7:23-31 [FREE Full text] [doi: 10.1016/j.invent.2016.12.003] [Medline: 28286739]
39. Choe E, Lee N, Lee B, Pratt W, Kientz J. Understanding quantified-selfers' practices in collectingexploring personal data. 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 26 April - 1 May 2014; Toronto, Canada. [doi: 10.1145/2556288.2557372]
40. Yuan S, Ma W, Kanthawala S, Peng W. Keep Using My Health Apps: Discover Users' Perception of Health and Fitness Apps with the UTAUT2 Model. Telemed J E Health 2015 Apr 28:735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
41. Bhuyan SS, Lu N, Chandak A, Kim H, Wyant D, Bhatt J, et al. Use of Mobile Health Applications for Health-Seeking Behavior Among US Adults. J Med Syst 2016 Jun;40(6):153. [doi: 10.1007/s10916-016-0492-7] [Medline: 27147516]
42. Oh J, Lee U. Exploring UX issues in Quantified Self technologies. 2015 Presented at: Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU); 20-22 Jan. 2015; Hakodate, Japan. [doi: 10.1109/icmu.2015.7061028]
43. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
44. Cordeiro F, Bales E, Cherry E, Fogarty J. Rethinking the mobile food journal: Exploring opportunities for lightweight photo-based capture. 2015 Presented at: 33rd Annual ACM Conference on Human Factors in Computing Systems; 18-23 April 2015; Seoul, South Korea. [doi: 10.1145/2702123.2702154]
45. Chan A, Kow R, Cheng JK. Adolescents' Perceptions on Smartphone Applications (Apps) for Health Management. Journal of Mobile Technology in Medicine 2017 Aug;6(2):47-55. [doi: 10.7309/jmtm.6.2.6]
46. Aljuraiban GS. Use of Weight-Management Mobile Phone Apps in Saudi Arabia: A Web-Based Survey. JMIR Mhealth Uhealth 2019 Feb 22;7(2):e12692 [FREE Full text] [doi: 10.2196/12692] [Medline: 30794205]
47. 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] [Medline: 23598614]
48. Woldeyohannes H, Ngwenyama O. Factors influencing acceptance and continued use of mHealth apps. 2017 Presented at: International Conference on HCI in Business, Government, and Organizations; July 9-14, 2017; Vancouver, BC, Canada. [doi: 10.1007/978-3-319-58481-2_19]
49. Cordeiro F, Epstein D, Thomaz E, Bales E, Jagannathan A, Abowd G, et al. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. 2015 Presented at: 33rd Annual ACM Conference on Human Factors in Computing Systems; 18-23 April 2015; Seoul, South Korea. [doi: 10.1145/2702123.2702155]
50. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
51. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and Facilitators of the Use of Mobile Health Apps From a Security Perspective: Mixed-Methods Study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
52. Warnick JL, Pfammatter A, Champion K, Galluzzi T, Spring B. Perceptions of Health Behaviors and Mobile Health Applications in an Academically Elite College Population to Inform a Targeted Health Promotion Program. Int J Behav Med 2019 Apr;26(2):165-174. [doi: 10.1007/s12529-018-09767-y] [Medline: 30632092]

53. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and Physical Activity Apps: Perceived Effectiveness by App Users. JMIR Mhealth Uhealth 2016;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
54. Kwon M, Mun K, Lee JK, McLeod DM, D'Angelo J. Is mobile health all peer pressure? The influence of mass media exposure on the motivation to use mobile health apps. Convergence 2016 Apr 10;23(6):565-586. [doi: 10.1177/1354856516641065]
55. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
56. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
57. Ghelani DP, Moran LJ, Johnson C, Mousa A, Naderpoor N. Mobile Apps for Weight Management: A Review of the Latest Evidence to Inform Practice. Front. Endocrinol 2020 Jun 24;11:412. [doi: 10.3389/fendo.2020.00412]
58. 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] [Medline: 18624603]
59. 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;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
60. Turner-McGrievy GM, Wilcox S, Boutté A, Hutto BE, Singletary C, Muth ER, et al. The Dietary Intervention to Enhance Tracking with Mobile Devices (DIET Mobile) Study: A 6-Month Randomized Weight Loss Trial. Obesity 2017 Dec;25(8):1336-1342 [FREE Full text] [doi: 10.1002/oby.21889] [Medline: 28600833]
61. 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;119(9):1525-1532. [doi: 10.1016/j.jand.2019.03.013] [Medline: 31155474]
62. König LM, Renner B. Boosting healthy food choices by meal colour variety: results from two experiments and a just-in-time Ecological Momentary Intervention. BMC Public Health 2019 Jul 22;19(1):975 [FREE Full text] [doi: 10.1186/s12889-019-7306-z] [Medline: 31331299]
63. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. npj Digit. Med 2020 Jul 10;3(1):1-6. [doi: 10.1038/s41746-020-0306-7]
64. Weinstein ND, Lyon JE, Sandman PM, Cuite CL. Experimental evidence for stages of health behavior change: the precaution adoption process model applied to home radon testing. Health Psychol 1998 Sep;17(5):445-453. [doi: 10.1037//0278-6133.17.5.445] [Medline: 9776003]
65. Attig C, Franke T. I track, therefore I walk – Exploring the motivational costs of wearing activity trackers in actual users. International Journal of Human-Computer Studies 2019 Jul;127:211-224. [doi: 10.1016/j.ijhcs.2018.04.007]
66. Lally P, Gardner B. Promoting habit formation. Health Psychology Review 2013 May;7(sup1):S137-S158. [doi: 10.1080/17437199.2011.603640]
67. Stawarz K, Cox A, Blandford A. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. : ACM; 2015 Presented at: 33rd annual ACM conference on human factors in computing systems; 18-23 April 2015; Seoul, South Korea. [doi: 10.1145/2702123.2702230]
68. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen S, Hamoy KG, et al. Mobile Ecological Momentary Diet Assessment Methods for Behavioral Research: Systematic Review. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11170 [FREE Full text] [doi: 10.2196/11170] [Medline: 30459148]
69. Burke LE, Warziski M, Starrett T, Choo J, Music E, Sereika S, et al. Self-monitoring dietary intake: current and future practices. J Ren Nutr 2005 Jul;15(3):281-290. [Medline: 16007557]
70. Ziesemer K, König L, Boushey C, Villinger K, Wahl D, Butscher S, et al. Occurrence of and Reasons for "Missing Events" in Mobile Dietary Assessments: Results From Three Event-Based Ecological Momentary Assessment Studies. JMIR MHealth UHealth 2020;8(10):e15430. [doi: 10.2196/preprints.15430]
71. Thomas JG, Bond DS, Ryder BA, Leahey TM, Vithiananthan S, Roye GD, et al. Ecological momentary assessment of recommended postoperative eating and activity behaviors. Surg Obes Relat Dis 2011;7(2):206-212. [doi: 10.1016/j.soard.2010.10.007] [Medline: 21130703]
72. Boushey CJ, Spoden M, Zhu FM, Delp EJ, Kerr DA. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proc Nutr Soc 2016 Dec 12:1-12. [doi: 10.1017/S0029665116002913] [Medline: 27938425]
73. 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] [Medline: 19916637]
74. 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;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]

75. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2016 Dec 13:254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
76. 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] [Medline: 17592961]
77. Klusmann V, Musculus L, Sproesser G, Renner B. Fulfilled Emotional Outcome Expectancies Enable Successful Adoption and Maintenance of Physical Activity. Front Psychol 2016;6:1990 [FREE Full text] [doi: 10.3389/fpsyg.2015.01990] [Medline: 26779095]
78. Gellert P, Ziegelmann JP, Schwarzer R. Affective and health-related outcome expectancies for physical activity in older adults. Psychol Health 2012;27(7):816-828. [doi: 10.1080/08870446.2011.607236] [Medline: 21867397]
79. Wahl DR, Villinger K, König LM, Ziesemer K, Schupp HT, Renner B. Healthy food choices are happy food choices: Evidence from a real life sample using smartphone based assessments. Scientific Reports 2017 Dec 06;7(1):17069. [doi: 10.1038/s41598-017-17262-9] [Medline: 29213109]
80. Bandura A. Social Cognitive Theory: An Agentic Perspective. Annu. Rev. Psychol 2001 Feb;52(1):1-26. [doi: 10.1146/annurev.psych.52.1.1]
81. Schwarzer R. Modeling Health Behavior Change: How to Predict and Modify the Adoption and Maintenance of Health Behaviors. Applied Psychology 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
82. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of Steel: Are Exercise Apps Providing Consumers With Realistic Expectations? Health Educ Behav 2012 Sep 17;40(2):133-139. [doi: 10.1177/1090198112452126]
83. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
84. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-t]
85. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. Journal of Nutrition Education and Behavior 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006]
86. Kelley CCG, Neufeld JM, Musher-Eizenman DR. Drive for thinness and drive for muscularity: opposite ends of the continuum or separate constructs? Body Image 2010 Jan;7(1):74-77. [doi: 10.1016/j.bodyim.2009.09.008] [Medline: 19944659]
87. Petticrew M, Egan M, Thomson H, Hamilton V, Kunkler R, Roberts H. Publication bias in qualitative research: what becomes of qualitative research presented at conferences? Journal of Epidemiology & Community Health 2008;62(6):552-554. [doi: 10.1136/jech.2006.059394]
88. Franke T, Attig C, Wessel D. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. International Journal of Human–Computer Interaction 2018;35(6):456-467. [doi: 10.1080/10447318.2018.1456150]
89. DiClemente C, Carbonari J, Velasquez M. Alcoholism treatment mismatching from a process of change perspective. In: Alcohol abuse treatment. Totowa, NJ: The Humana Press; 1992:115-142.
90. Graham AK, Lattie EG, Mohr DC. Experimental Therapeutics for Digital Mental Health. JAMA Psychiatry 2019 Aug 21;76(12):1223-1224. [doi: 10.1001/jamapsychiatry.2019.2075] [Medline: 31433448]
91. Baretta D, Perski O, Steca P. Exploring Users' Experiences of the Uptake and Adoption of Physical Activity Apps: Longitudinal Qualitative Study. JMIR Mhealth Uhealth 2019;7(2):e11636 [FREE Full text] [doi: 10.2196/11636] [Medline: 30735143]
92. European Commission. eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century. Brussels: European Commission; 2012.
93. European Commission. eHealth: Digital health and care. 2017. URL: https://ec.europa.eu/health/ehealth/policy_en [accessed 2021-06-14]

Abbreviations

BCT: behavior change theory
DG: design guideline

Edited by G Strudwick; submitted 12.05.20; peer-reviewed by A Rhodes, A Heidel; comments to author 03.06.20; revised version received 11.01.21; accepted 01.04.21; published 19.06.21.

Please cite as:
König LM, Attig C, Franke T, Renner B
Barriers to and Facilitators for Using Nutrition Apps: Systematic Review and Conceptual Framework
JMIR Mhealth Uhealth 2021;9(6):e20037
URL: https://mhealth.jmir.org/2021/6/e20037/
doi:10.2196/20037
PMID:

Original Paper

App-Based Versus Standard Six-Minute Walk Test in Pulmonary Hypertension: Mixed Methods Study

Dario Salvi[1], PhD; Emma Poffley[2], MSc; Lionel Tarassenko[1], DPhil, FMedSci; Elizabeth Orchard[2], MBBS, MA

[1]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom

[2]Department of Cardiology, Oxford University National Health Service Foundation Trust, Oxford, United Kingdom

Corresponding Author:
Dario Salvi, PhD
Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford
Old Road Campus Research Building
Oxford
United Kingdom
Phone: 44 1865 617675
Email: dario.salvi.work@gmail.com

Abstract

Background: Pulmonary arterial hypertension (PAH) is a chronic disease of the pulmonary vasculature that can lead to heart failure and premature death. Assessment of patients with PAH includes performing a 6-minute walk test (6MWT) in clinics. We developed a smartphone app to compute the walked distance (6MWD) indoors, by counting U-turns, and outdoors, by using satellite positioning.

Objective: The goal of the research was to assess (1) accuracy of the indoor 6MWTs in clinical settings, (2) validity and test-retest reliability of outdoor 6MWTs in the community, (3) compliance, usability, and acceptance of the app, and (4) feasibility of pulse oximetry during 6MWTs.

Methods: We tested the app on 30 PAH patients over 6 months. Patients were asked to perform 3 conventional 6MWTs in clinic while using the app in the indoor mode and one or more app-based 6MWTs in outdoor mode in the community per month.

Results: Bland-Altman analysis of 70 pairs of conventional versus app-based indoor 6MWDs suggests that the app is sometimes inaccurate (14.6 m mean difference, lower and upper limit of agreement: –133.35 m to 162.55 m). The comparison of 69 pairs of conventional 6MWDs and community-based outdoor 6MWDs within 7 days shows that community tests are strongly related to those performed in clinic (correlation 0.89), but the interpretation of the distance should consider that differences above the clinically significant threshold are not uncommon. Analysis of 89 pairs of outdoor tests performed by the same patient within 7 days shows that community-based tests are repeatable (intraclass correlation 0.91, standard error of measurement 36.97 m, mean coefficient of variation 12.45%). Questionnaires and semistructured interviews indicate that the app is usable and well accepted, but motivation to use it could be affected if the data are not used for clinical decision, which may explain low compliance in 52% of our cohort. Analysis of pulse oximetry data indicates that conventional pulse oximeters are unreliable if used during a walk.

Conclusions: App-based outdoor 6MWTs in community settings are valid, repeatable, and well accepted by patients. More studies would be needed to assess the benefits of using the app in clinical practice.

Trial Registration: ClinicalTrials.gov NCT04633538; https://clinicaltrials.gov/ct2/show/NCT04633538

(JMIR Mhealth Uhealth 2021;9(6):e22748) doi:10.2196/22748

KEYWORDS

cardiology; exercise test; pulmonary hypertension; mobile apps; GPS

Introduction

Pulmonary arterial hypertension (PAH) is a progressive illness that can be a severe life-limiting condition if not diagnosed early or left untreated [1]. It is a chronic disease of the pulmonary vasculature, with vascular proliferation and remodeling of the small pulmonary arteries leading to a progressive increase in pulmonary vascular resistance. This ultimately results in right

heart failure and premature death [2]. The predominant symptom of PAH is dyspnea on exertion, with a decrease in exercise capacity. PAH is an uncommon condition, affecting about 6000 patients in the United Kingdom [3].

The 6-minute walk test (6MWT) is a standard method for measuring exercise capacity in patients with cardiopulmonary disease such as PAH. The 6MWT measures how far a patient can walk in 6 minutes [4]. Walking is an activity performed every day by most patients except for those most severely limited. By assessing patients' ability to exercise, the 6MWT provides a global assessment of respiratory, cardiovascular, neuromuscular, and cognitive function. The 6MWT does not differentiate what limits the patient nor does it assess maximal exercise capacity. Instead, the 6MWT allows the patient to exercise at a daily functional level and is a useful tool for assessing severity of disease, and increasing walk distance correlates with a subjective improvement in dyspnea [5].

In PAH, the 6MWT is used to assess patients' risk of death, with a walked distance of >440 m being associated with a low risk (<5% a year estimated risk of mortality) and <165 m being a high risk (>10% risk of mortality) [6]. An increase in 6MWT distance of more than 42 m is considered a clinically significant improvement [7]. Furthermore, change in 6MWT distance correlates with VO_2 max, New York Heart Association class, and mortality in PAH patients, providing an objective assessment of disease progression, prognosis, and response to treatment [8]. It is a universally accepted test as it is safe and easily performed by the patient.

The 6MWT has become the primary end point for many trials and as a surrogate for invalidated survival outcome for all placebo-controlled trials of PAH therapy [9]. It is used by regulatory bodies to determine whether a treatment should be approved. The test is usually performed in the hospital, by having the patient walk along a hospital corridor, where many factors on the day of the test can affect patient performance, including tiredness, lack of familiarity with the environment, or anxiety. Two physiologists are typically required to monitor the test to measure distance walked and oxygen saturations by pulse oximetry and record symptoms felt during the test. This justifies the need for a community-based approach that would allow a more accurate and frequent measure of the day-to-day function of PAH patients while reducing employed resources.

Community-based 6MWTs have been performed in chronic stroke patients using a GPS tracker [10] and in heart failure patients using accelerometers and step counters [11]. A smartphone-based method was used by Ata et al [12] for assessing patients with peripheral artery disease, but it relied on the embedded distance measurement of the iPhone and proved to have low accuracy. More promising results were achieved using a custom smartphone app in congestive heart failure and pulmonary hypertension participants [13] or our app, named SMWT, which demonstrated high accuracy and user acceptance in lab tests [14].

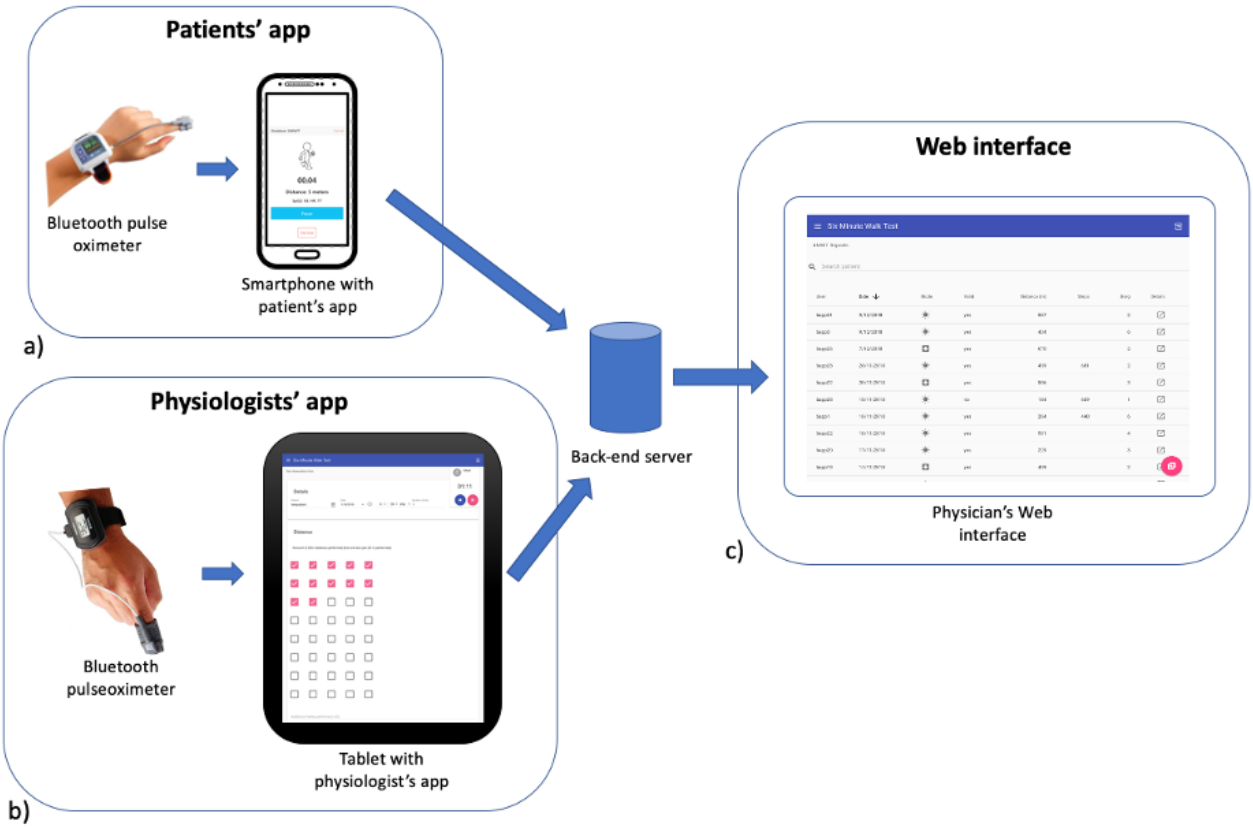
In this study, we aimed to test our app on 30 patients with pulmonary hypertension. The aim of the study was to assess if patients were able and willing to use the app, compare the 6MWT distance (6MWD) measured by the app with the reference measured in the clinic by physiologists, and explore the feasibility of pulse oximetry during app-based 6MWT.

Methods

Approach

The SMWT system is described in detail in Salvi et al [14]; nonetheless, a summary of its architecture and functionalities is provided here for reference. The system comprises a mobile phone app for patients, tablet app used by physiologists, and server (Figure 1). The patient app can work in two modalities: indoor, where inertial sensors are used to measure the number of U-turns the patient performs while walking on a straight walkway, and outdoor, where the positioning system, like the GPS, is used to track the user and compute the walked distance. The app also allows users to connect to a Bluetooth pulse oximeter, supporting both the $WristOx_2$ 3150 (Nonin) and the PC-68B (Shenzhen Creative Industry Co Ltd) medical sensors.

Figure 1. Architecture of the SMWT system: (a) patients' app, (b) physiologists' app, and (c) web interface for physicians.



The server collects the 6MWT data from the app and provides a web interface to review the data. Physiologists can also add conventional 6MWT information like total walked distance and symptoms on the patient's web page. The physiologist app offers the same interface as the website, with the added possibility to collect live data from the WristOx$_2$ 3150 pulse oximeter.

We defined a study protocol with the aim of demonstrating that patients are able and willing to use the app for the 6MWT. The main outcomes were percentage of participants who performed at least 1 community 6MWT and percentage of patients who provided at least 1 community 6MWT per month during the study period. Secondary outcomes were as follows:

- Comparison of the indoor app-based 6MWD against simultaneous clinic 6MWD (statistics of differences, Bland-Altman analysis)
- Comparison between community-based outdoor 6MWD and clinic 6MWD within 7 days (statistics of differences, Bland-Altman analysis)
- Assessment of test-retest reliability of the community-based 6MWT within 7 days (statistics of differences in consecutive 6MWDs, two-way random effects model, single measures, absolute agreement intraclass correlation [15], standard error of measurement computed as standard deviation of 6MWD [16], mean intrasubject coefficient of variation)
- Assessment of usability and acceptance (Mobile Application Rating Scale–user version [17] and mobile health acceptance [18] questionnaires and semistructured interviews)
- Feasibility of pulse oximetry during app-based 6MWT (Bland-Altman analysis of pulse oximetry collected by two sensors simultaneously)

Selection Criteria

Inclusion criteria were willing and able to give informed consent for participation in the study, male or female aged 18 years and older, diagnosed with PAH but able to undertake a 6MWT off-oxygen, PAH group 1 (regular PAH) or 4 (pulmonary hypertension due to blood clots in the lungs), own a compatible smartphone (Android or iPhone) and able to use it. Exclusion criteria were use of long-term oxygen therapy; cognitive impairments; rheumatological diseases that limit the measurement of finger oxygen saturations; PAH groups 2 (pulmonary hypertension due to left heart disease), 3 (pulmonary hypertension due to lung disease), or 5 (blood and other rare disorders that lead to pulmonary hypertension); cannot use a smartphone; or pregnancy.

We focused on patients in PAH groups 1 and 4 [19] because they have treatable forms of PAH and need to perform frequent 6MWTs in order to titrate their medication regime for pulmonary hypertension and assess their response to it. Patients with cognitive impairments were excluded because they would have struggled with using the app. Patients who require oxygen to mobilize would have needed to walk and carry their oxygen supply, and this would have affected the distance that they could walk outside. Pregnancy was excluded so as to not introduce confounding factors.

Patients were recruited during regular PAH clinics. Each patient participated for about 6 months. They performed 3 conventional

6MWTs at the clinic, one on the day of recruitment, one after 3 months and one after 6 months. The clinic 6MWT was performed while having the patient's phone running the app in indoor mode. Patients were additionally asked to perform an app-supported outdoor 6MWT in their community within 3 days of the clinic.

Between clinics, patients were invited to perform one community-based 6MWT preferably every 15 days and not less than once per month. To improve compliance, patients who did not provide a community 6MWT within 1 month were reminded either by email or phone call. At the last clinic 6MWT, participants were given usability and acceptance questionnaires [17,18] to be completed either online or on paper. A number of patients were also interviewed depending on their and the staff availability. Paper-based answers were digitized by one team member. Interviews were performed by one team member and audiorecorded with explicit signed consent. The questions were based on a technology acceptance model [20]. Audio was transcribed by one team member who also analyzed the transcriptions.

In order to guarantee quality of the collected data, both in-clinic and community-based measurements were visually assessed using the web interface by an engineer and a physiologist on a weekly basis. Tests that showed evident artefacts or that patients did not comply with recommendations about how to perform the test (ie, when the walked path showed more than 3 very narrow curves or if the variation of altitude over the entire path was more than 20 m), were flagged as invalid and not included in the analysis (examples of invalid tests are provided in Multimedia Appendix 1). The data were extracted from the database used in the system (ArangoDB) and analyzed with the R programming language (R Foundation for Statistical Computing).

The study protocol was approved by the National Health Service Health Research Authority (protocol reference number: 17/WM/0355) and registered at Oxford University Hospitals National Health Service Foundation Trust. The study was also registered at ClinicalTrials.gov [NCT04633538]. Written informed consent was obtained from all participants.

Results

Participants and Tests

We approached 33 patients; of these, 30 were eligible and consented to participate. Five patients withdrew from the study (3 were lost to follow-up, 1 died, and 1 considered the study not to be relevant any longer). Of the patients who started, 37% (11/30) were male and 63% (19/30) were female. Their mean age was 50 (SD 16.55; range 20-76) years. The first patient was enrolled on January 25, 2018, and the latest patient completed the study on September 7, 2019. Patients were actively enrolled in the study for a mean of 244.23 (SD 96.16; range 147-590) days.

In total, we received 455 test reports, of which 429 were considered valid according to our data quality procedures. Each patient performed 3 6MWTs in the clinic, 3 months apart from each other, using the app in indoor mode while the physiologist reported their observations (distance, oxygen saturation, Borg scale, symptoms) using the dedicated website or app. We received 81 test reports from the physiologist and 71 app-based tests in indoor mode. Between clinic visits, patients were asked to perform tests in the community using the app in outdoor mode. We received 277 such test results sent by 12 Android phones and 29 iPhones (9 patients used more than one phone during their participation in the study).

Accuracy of the App 6MWT in Indoor Mode

We compared 70 pairs of 6MWD, for which one was measured by a physiologist during the clinic 6MWT and the other was simultaneously computed by the app in indoor mode. The tests belong to 29 patients. To compare the two distances, we report the statistics in Table 1 and Bland-Altman plot in Figure 2.

To understand if the accuracy was affected by a systematic bias created by some phones (eg, because of different sensor characteristics), we plotted the absolute difference of the distance between the app-computed 6MWD and the reference value for each phone and each user in Figure 3.

Table 1. Differences between a 6-minute walk test distance (6MWD) measured by physiologist and simultaneous 6MWD measured by the app in indoor mode.

Characteristic	Value
Difference (m), mean (SD)	14.6 (75.48)
Difference (% of total length), mean (SD)	1.92 (18.6)
Difference (m), median (min-max)	2.65 (–209.8 to 339.64)
Difference (% of total length), median (min-max)	0.65 (–58.93 to 81.64)
Correlation (*P* value)	0.83 (<.001)
Percentage of difference below 42 m (%)	83
Lower and upper limits of agreement (m)	–133.35 to 162.55

Figure 2. Bland-Altman plot of the difference between the 6-minute walk test distance as measured by the app in indoor mode and as observed by the physiologist during a 6-minute walk test in clinic. 6MWD: 6-minute walk test distance.

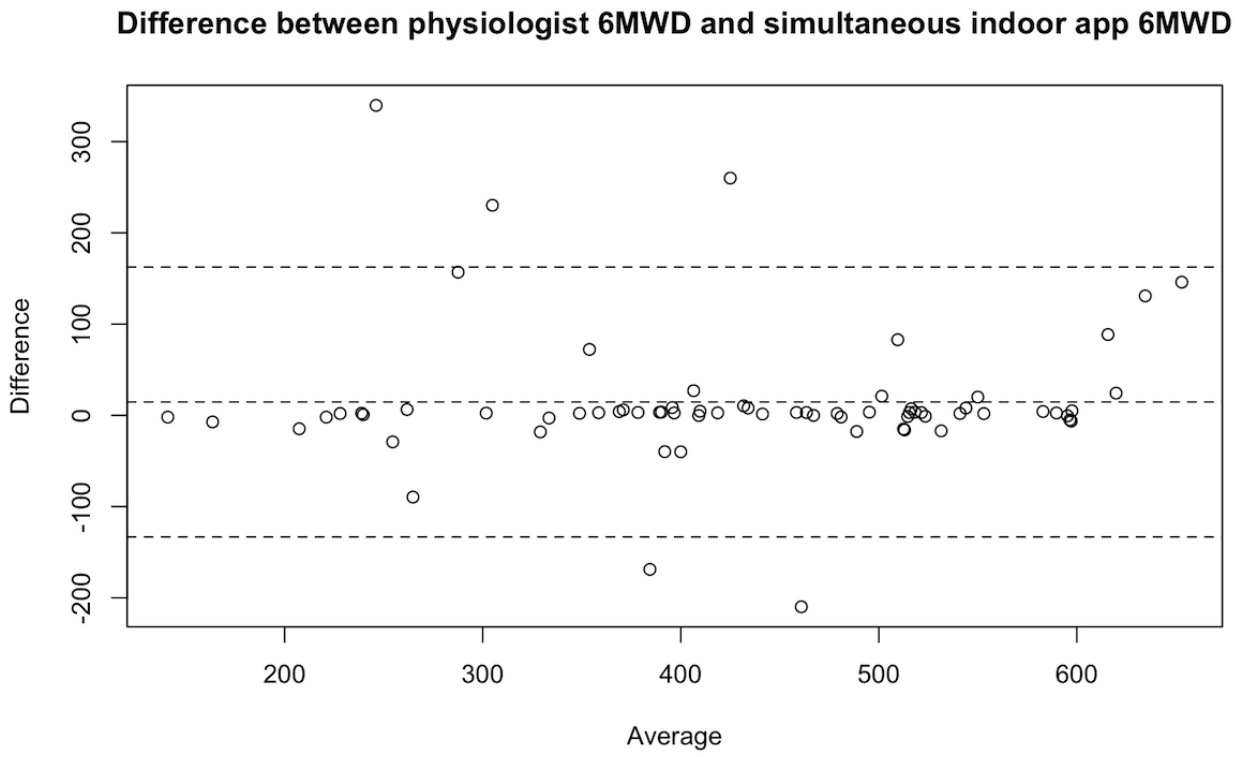
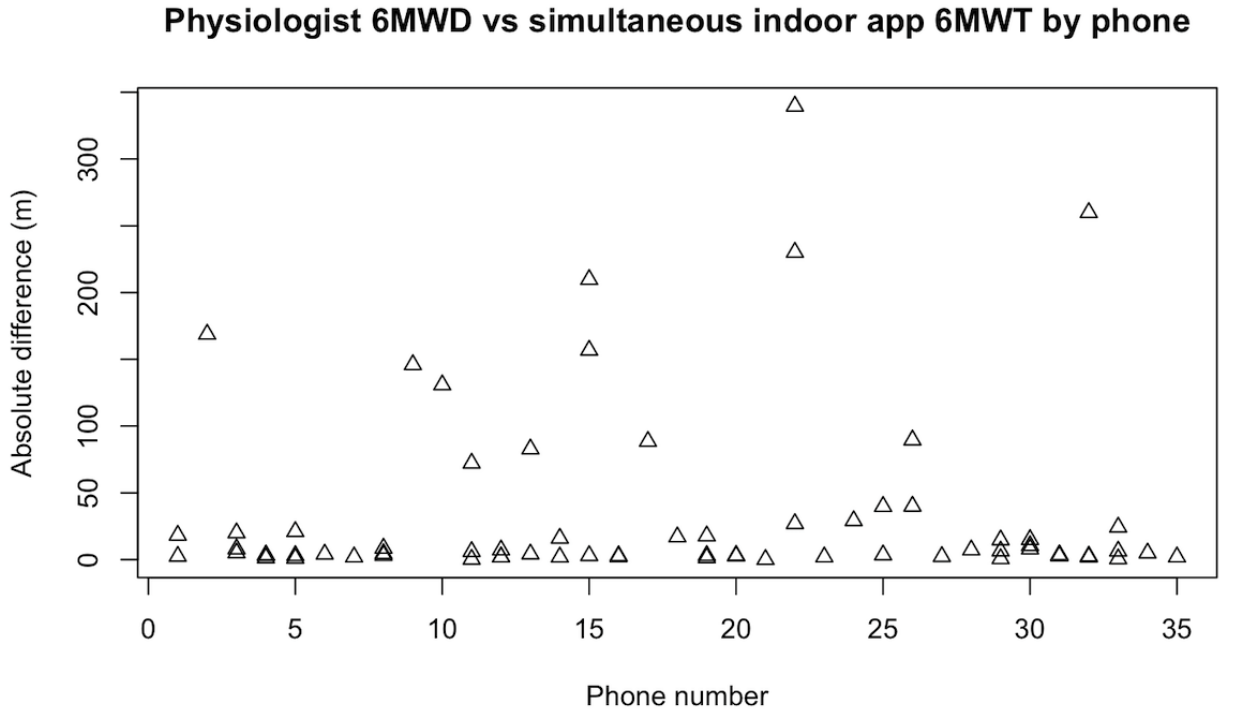


Figure 3. Absolute difference of the 6-minute walk test distance as measured by the physiologist and as measured by the mobile phone app in indoor mode. 6MWD: 6-minute walk test distance; 6MWT: 6-minute walk test.



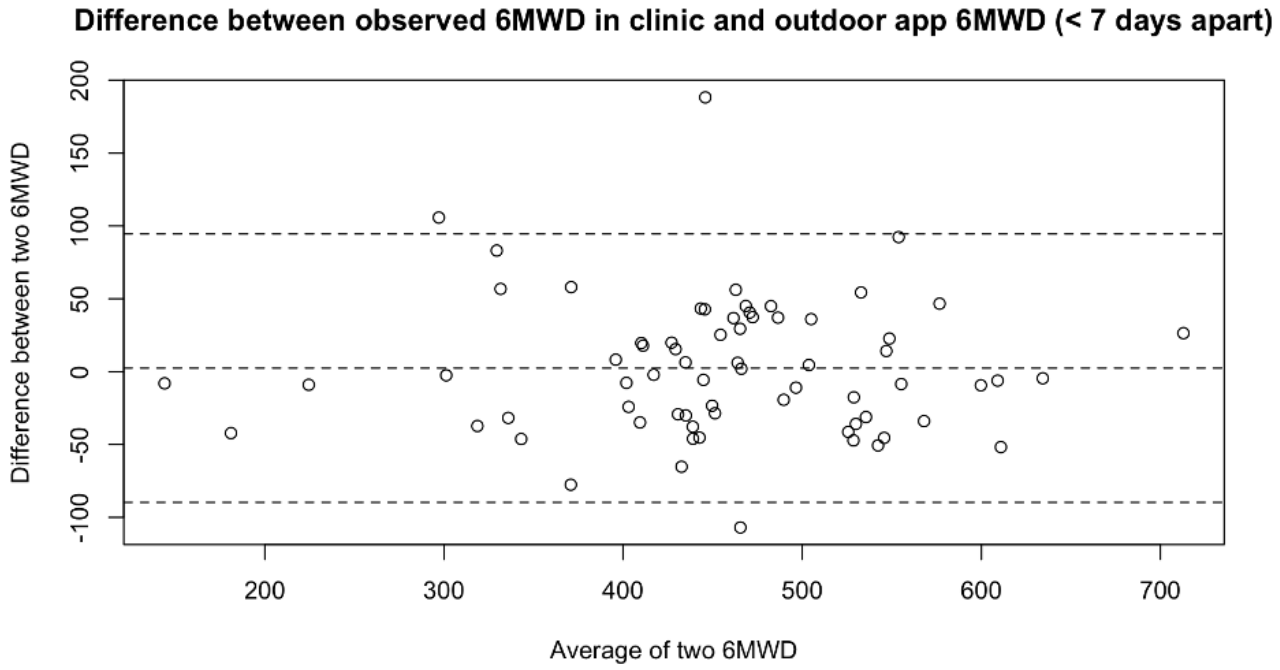
Validity of Community-Based Outdoor 6MWD

We compared pairs of 6MWD estimates where one was the result of an observation done by a physiologist in a clinic and the other was measured by the app during an outdoor test performed within 7 days of the clinic test. We analyzed 69 pairs belonging to 22 patients. Statistics for the differences and Bland-Altman plots are provided in Table 2 and Figure 4, respectively.

Table 2. Differences between 69 pairs containing a 6-minute walk test distance (6MWD) estimated by a physiologist during a clinic and a 6MWD measured by the app in outdoor mode within 7 days of the clinic test.

Characteristic	Value
Difference (m), mean (SD)	2.45 (47.0)
Difference (% of total length), mean (SD)	–0.18 (11.04)
Difference (m), median (min-max)	–4.55 (–106.91 to 188.24)
Difference (% of total length), median (min-max)	–0.83 (–26.35 to 34.86)
Correlation, (*P* value)	0.89 (<.001)
Percentages of difference below 42 m (%)	65
Lower and upper limits of agreement (m)	–89.67 to 94.57

Figure 4. Bland-Altman plot of the difference between the 6-minute walk test distance as observed by the physiologist during a 6-minute walk test in clinic and as measured by the app in outdoor mode within 7 days of the clinic test. 6MWD: 6-minute walk test distance.



We visually inspected the GPS traces of the 3 cases where the absolute difference was above 100 m and observed that, in 2 of those cases, the patient walked in small circles or performed numerous U-turns, which affects the accuracy of the distance estimation algorithm [14].

Test-Retest Reliability of Community-Based 6MWT

The 6MWD estimates provided at the end of community-based outdoor 6MWTs were stable over time (examples shown in Figure 5). To quantify this, we analyzed 89 pairs of outdoor tests performed by the same patient no more than 7 days apart. These pairs belonged to 10 patients in total. Statistics of the differences are reported in Table 3 together with intraclass correlation coefficient and standard error of measurement.

Figure 5. Outdoor, community-based 6-minute walk test distances over time of the top 5 patients who contributed with the most 6-minute walk tests. 6MWD: 6-minute walk test distance.

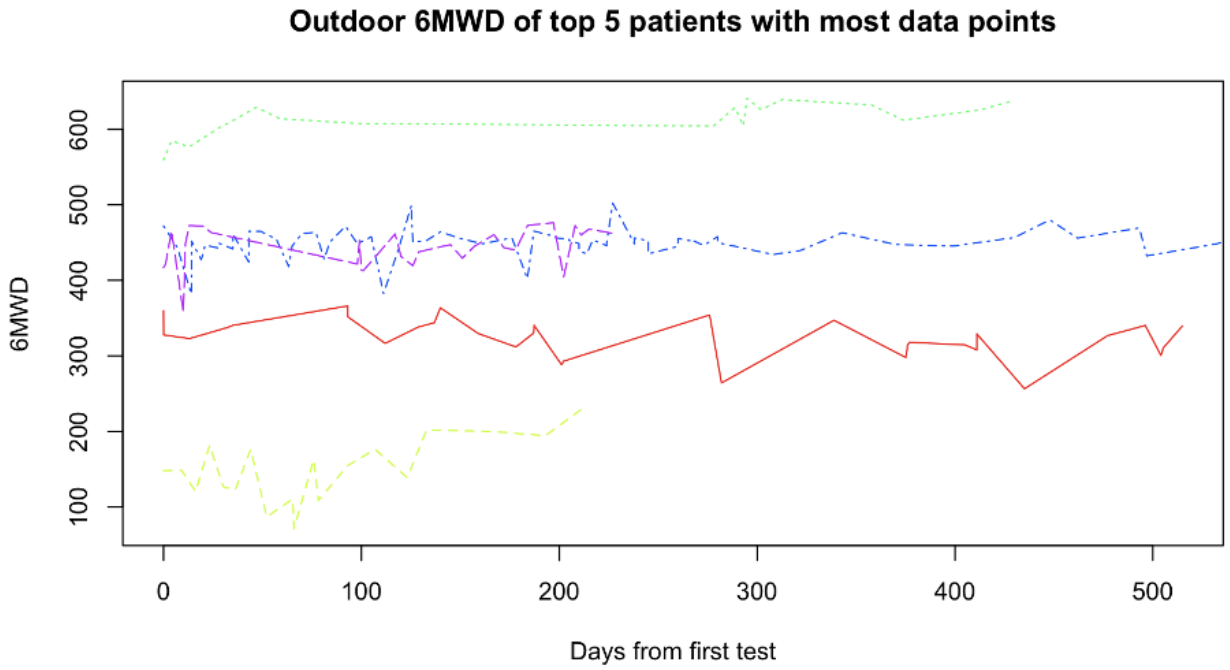


Table 3. Differences between the 6-minute walk test distances (6MWDs) of 89 pairs of community-based 6MWDs measured by the app in outdoor mode no more than 7 days apart.

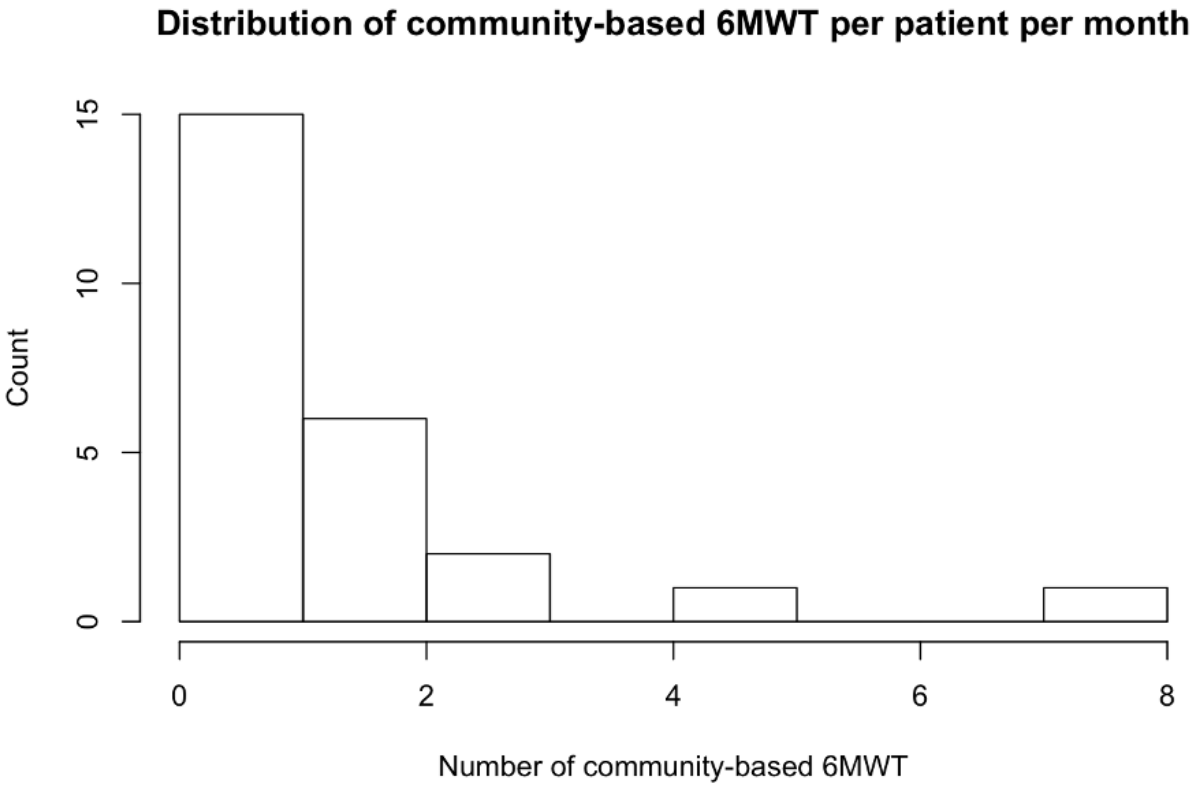
Characteristic	Value
Difference (m), mean (SD)	1.80 (36.97)
Difference (% of total length), mean (SD)	0.72 (10.13)
Difference (m), median (min-max)	0.95 (–91.69 to 107.90)
Difference (% of total length), median (min-max)	0.20 (–25.64 to 35.26)
Correlation, (*P* value)	0.93 (<.001)
Percentages of difference below 42 m (%)	80
Intraclass correlation	0.91
Standard error of measurement (m)	36.97
Coefficient of variation (%), mean	12.45

We manually inspected the only pair of tests that presented a difference of 6MWD higher than 100 m and it appeared that in one of the tests, the user walked in a narrow path with U-turns.

Compliance, Usability, and Acceptance

All patients who completed the study sent the results of at least one community-based 6MWT during the study, and 48% (12/25) of those who completed the study provided at least 1 6MWT result per month. The distribution of the number of community-based 6MWTs performed per patient per month is shown in Figure 6.

Figure 6. Histogram of the number of community-based 6-minute walk tests per patient per month. 6MWT: 6-minute walk test.



A total of 18 patients completed the usability and acceptance questionnaire. Likert scales to rate usability from 1 (very negative opinion) to 5 (very positive opinion) were all marked 3 and above. Also, the perceived impact of the app scored highly, with all average scores above 3. The answers provided to the acceptance questionnaire were above 3 on average except the question related to the perceived vulnerability construct ("I do not think my doctor understands how I feel doing my daily activities"). The exact wording of the questions and statistics of the answers are provided in Multimedia Appendix 2.

The questionnaire concluded with open questions about negative and positive aspects of the app and possible improvements. Answers with similar topics were grouped by one researcher and summarized.

Patients were happy with the app because it is easy to use (9/18), allows patients to perform the test within their community instead of having to go to the hospital (4/18), indicates the fitness level and oxygen saturation (4/18), and invites patients to walk regularly (1/18).

Reasons for concern were that it sometimes takes too long or fails to send the data (4/18), some elements of the user interface (eg, how to cancel a test) are not clear (2/18), patients forget to use it (1/18), log-in procedure is complicated (1/18), if a test fails it must be repeated (1/18), user interface could be more colorful and less medical (1/18), it does not accurately compute steps (1/18), pulse oximeter does not work correctly in cold weather (1/18), and some patients find it hard to walk for 6 minutes (1/18).

As for improvements, patients suggested providing charts about pulse oximetry on the app (3/18), integrate a wearable device (1/18), improve the feedback on incorrect log-in (1/18), make the detection of GPS satellites faster (1/18), improve user interface (1/18), and give the possibility of shortening the time of the test (1/18).

We interviewed 12 patients with a semistructured format (questions in Multimedia Appendix 3) on their last day in the study. Both notes taken during the interviews and transcripts were later analyzed by one researcher who grouped similar topics and summarized the feedback.

Motivating factors for using the app were self-monitor health status (7/12), pushing oneself to walk more (2/12), wanting to help research (2/12) and the feeling of being monitored by caregivers (1/12). In order to increase use, 1 patient suggested including reminders. The app was seen as useful (7/12), easy to use (3/12) and allowing better 6MWT than in the clinic because of not having to turn around (1/12).

Factors that contributed to not using the app were lack of time (3/12), weather (either too cold or too warm; 3/12), laziness (1/12), concerns about privacy (ie, not wanting to be seen walking holding the phone and pulse oximeter; 1/12), or lack of interest because of improved health conditions (1/12). Technical issues were reported about the app being too slow sending data (5/12), short battery life of the pulse oximeter (2/12), or inaccuracy compared to distance reported on Google Maps (1/12). Patients explicitly reported the app or technology not being an issue for motivation (1/12) nor concerns about safety (8/12) or data protection (7/12).

Most patients felt the app made them feel more connected with the caregiver who could review the data (10/12), and some would even trust it to substitute their periodic appointment with the doctor (5/12). In terms of suggestions, patients thought they wanted to have more comprehensive data about past tests (eg, pulse oximetry; 3/12), integration with fitness trackers and apps (3/12), reminders (2/12), and a simpler Borg scale (1/12). More than half of the patients (7/12) indicated they would be willing to use the app after the end of the study.

Reliability of Pulse Oximetry

While monitoring data quality in the study, we observed that in 33.6% (130/387) of tests where pulse oximetry was recorded, the pulse oximeter presented a peripheral oxygen saturation (SpO_2) value higher 3 minutes after the start of the test than at baseline. Likewise, 28.2% (109/387) of tests showed a decrease in the heart rate (HR) after 3 minutes with respect to the baseline. In order to explain this unexpected behavior, we analyzed the traces of SpO_2 and HR for some of these tests (an example of such a test is shown in Figure 7). We hypothesized that this behavior was due to how the pulse oximeters were used; therefore, during clinic 6MWTs, we began collecting pulse oximetry from both the patient app, which uses the PC-68B, and the physiologist app, which can connect to the WristOx$_2$ 3150. In addition, we asked one patient to run the app on two independent phones when doing community-based 6MWTs, each phone using one of the two types of pulse oximeters.

With this strategy, we retrieved pairs of pulse oximetry samples (SpO_2 and HR) from 38 tests. Samples with the shortest time differences (5 seconds or less) were matched and interpolated, thereby obtaining 19,279 matched samples to be compared. Figure 8 shows the Bland-Altman plot. The lower and upper limits of agreement are –8.72% and 9.73% for SpO_2 and –33.87 and 26.97 bpm for HR.

Figure 7. SpO_2 and heart rate for a test where the SpO_2 increases during exertion and the heart rate decreases. SpO_2: peripheral oxygen saturation.

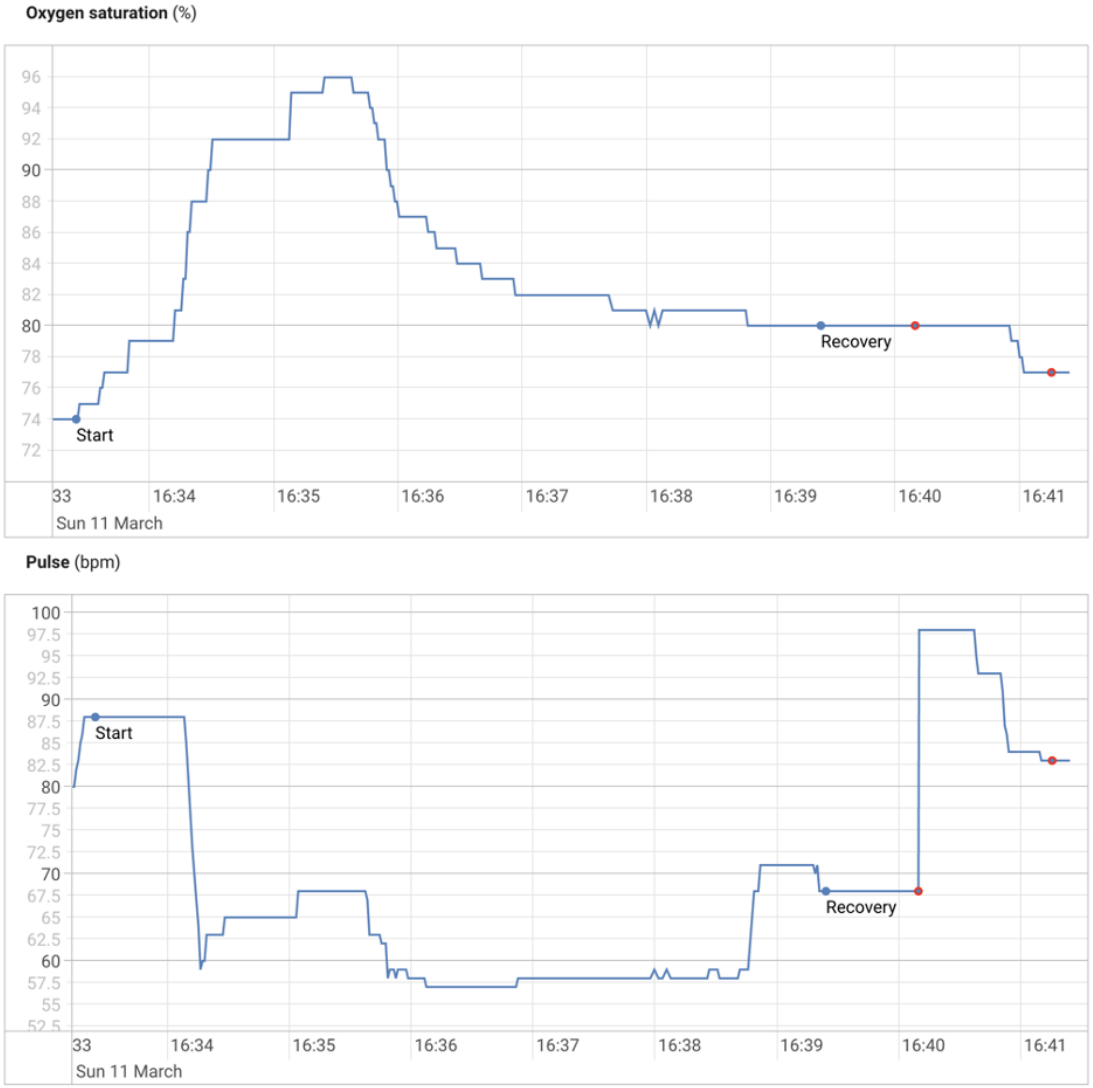
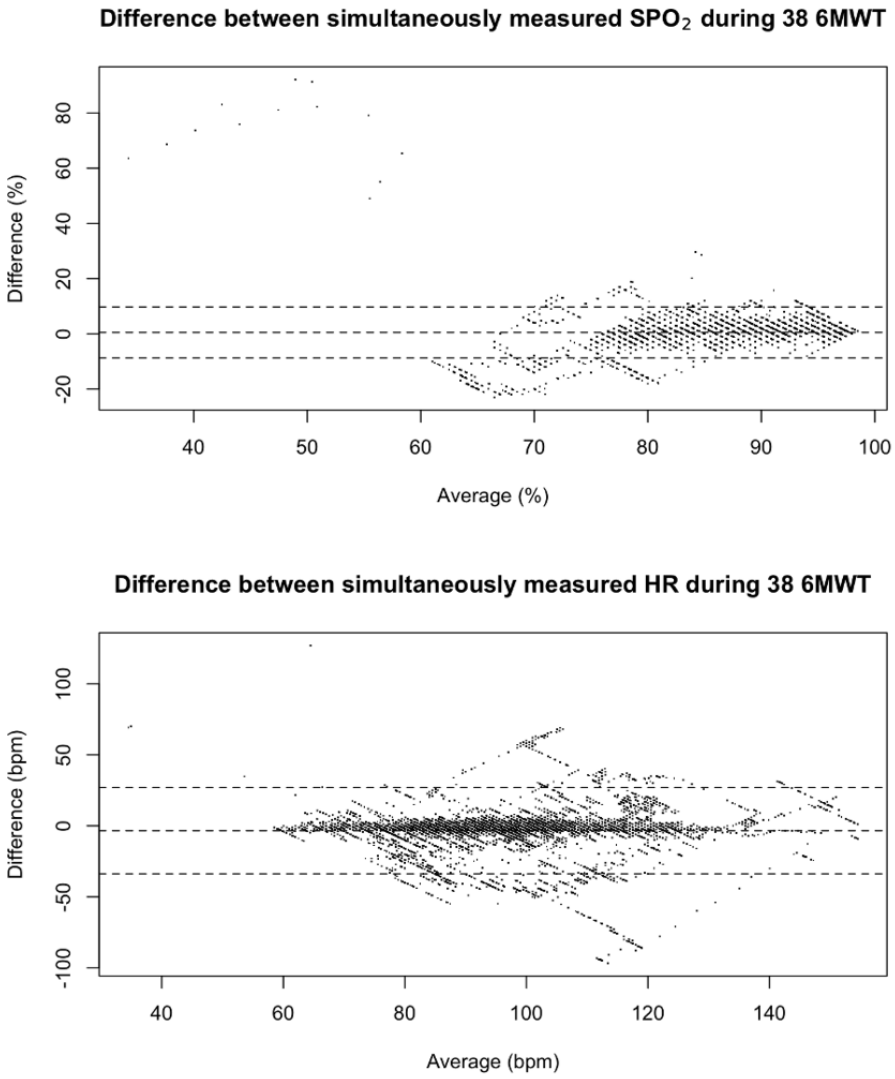


Figure 8. Bland-Altman plots of the differences between 19,356 matched samples of SpO_2 and heart rate values measured simultaneously by two different pulse oximeters during 38 6-minute walk tests (6MWTs). SpO_2: peripheral oxygen saturation.



Discussion

Accuracy of the App 6MWT in Indoor Mode

Compared to the lab tests of the same app reported in Salvi et al [14], the tests with patients' phones showed higher variability in the differences between the distance measured by the physiologist and the distance measured by the phone (–2.01 [SD 7.84] m in lab tests vs 14.6 [SD 75.48] m in our study). The distribution of error by phone does not suggest any systematic bias in some phones; therefore, the inaccuracy must be due to specific test conditions that are not well tolerated by the distance estimation algorithm.

The Bland-Altman analysis shows that the limits of agreement (–133.35 and 162.55 m) are above the clinically significant threshold of 42 m; therefore, the app-based indoor measurement cannot be considered equivalent to the physiologist observation. Nonetheless, the analysis also shows that in 83% of the measurements the difference was below that threshold, which suggests that the algorithm fails sporadically but when it does, it introduces a high error, thus skewing the statistics.

Validity of Community-Based Outdoor 6MWD

Our results show that the community-based, outdoor 6MWT correlated highly with the 6MWT performed in the clinic. This is similar to what is reported by Brooks et al [13] where a similar app is used (our 0.89 correlation coefficient vs their 0.88), although, in Brooks et al [13], a lower standard deviation of the difference is reported (26 m vs our 47 m).

While the high correlation confirms the general validity of the approach, Bland-Altman analysis shows that high differences can exist between clinic- and community-based measurements. These must be taken into account when interpreting the measured 6MWD, especially because the standard deviation of the difference was 47 m, which is above the clinically significant threshold. These differences can be explained by a combination of three factors: (1) inaccuracy of the technology: lab tests show that this amounts to up to 20 m of the standard deviation [14]; (2) variations in 6MWD when the test is performed on different days: analysis of conventional 6MWT in PAH shows a standard deviation of the difference of about 20 m [7,21]; and (3) some tests performed without following the instructions may have

passed through our data quality assurance measures: the algorithm is known to lose accuracy when the test is performed in confined spaces and/or in the presence of many curves, and this was confirmed at least in 2 of the 3 cases where the difference was above 100 m.

Test-Retest Reliability of Community-Based 6MWT

Consecutive tests performed within 7 days showed a high correlation (0.93). Compared to a similar app [13], we obtained a higher coefficient of variation (our 12.45% vs their 4.6%). Compared to the test-retest reliability of the GPS-method reported by Wevers et al [10], we obtained a similarly high intraclass correlation (our 0.91 vs their 0.96) but a higher standard error of measurement (our 36.97 m vs their 18.1 m). The standard deviation of the measurement error of the app as calculated in lab settings, 18.56 m [14], indicates that part of the standard error of measurement within consecutive measurements can be justified by the inaccuracy of the instrument. Nonetheless, it has to be noted that this value is lower than the clinically significant threshold of 42 m, as about 80% of the differences were below that threshold.

Compliance, Usability, and Acceptance

While 100% of patients were able to use the app at least once during the 6-month duration of the study, only 48% followed the recommendation to perform 1 or more community-based 6MWTs per month. Low compliance is a known issue in telehealth studies [22]. As interviews revealed, low compliance in our cohort was due to lack of time or motivation or patients forgetting to use the app. It should be observed, however, that compliance varied significantly from patients who performed only one community-based test in the whole study to patients who performed more than 7 tests per month.

A possible explanation for low motivation in some patients could be the low perceived vulnerability (ie, patients feeling that they do not need further monitoring/follow-up in addition to regular clinics). This is particularly relevant within the context of this study as patients were informed that the study investigator (a cardiologist) would not use the collected data for any clinical decision making. Other external factors like weather conditions or not wanting to be seen using the app and the pulse oximeter may have also contributed.

The results of the questionnaire and the interviews show that the app was easy to use and well accepted. This is in agreement with other similar apps, like the one reported in Brooks et al [13]. In addition, the app was perceived as safe, secure (from the data protection perspective), and useful. Patients particularly appreciated the knowledge of their health being monitored either by themselves or the clinicians. A few usability and technical issues related to lack of control for some parts of the test (eg, difficulty in restarting the test) or data transmission (too slow and lacking feedback about connection status) were identified. However, these technical issues were not reported as a barrier for the use of the app.

In terms of improvements, patients suggested including reminders of scheduled tests, which could also boost compliance, integrating the app with other fitness apps and wearables, and providing more detailed data, especially about past tests.

Reliability of Pulse Oximetry

The inaccuracy of pulse oximetry during walking is known in the literature [23-25] and may be due to arterial flow being significantly affected by motion. The two pulse oximeters employed in our studies were each certified as medical devices, and they each showed decreasing HR and increasing SpO_2 at times during walking. Comparison of the data produced by the devices suggests that their values differed significantly (–8.72% and 9.73% for SpO_2 and –33.87 and 26.97 bpm for HR), which is above what would be justified by the confidence interval reported by the manufacturers (±3 for both SpO_2 and HR). We could not determine which of the devices was more accurate.

Principal Findings

The main findings of this study are that, while the app-based indoor 6MWTs cannot be considered equivalent to the conventional one, the outdoor, community-based 6MWT measurements are strongly correlated with those performed in clinic and are repeatable. These community-based tests can be considered as a valid complement to conventional 6MWTs especially because they can be performed more frequently and at the patient's convenience, but patients must be instructed to use the app correctly to avoid inaccuracies.

Results also show that even if the app has proven to be usable and well accepted, its usefulness should be made explicit to patients to increase their engagement. Additionally, the use of conventional pulse oximetry should not be relied on in a 6MWT, at least during the walking stage.

Implications for Future Research

Future versions of this system would benefit from technical improvements. The insufficient accuracy shown in the indoor scenario suggests that new algorithmic approaches should be explored to make the measurement more robust. Algorithms used for dead-reckoning of pedestrians may offer a valid strategy [26]. In addition, the algorithm used for the outdoor scenario is penalized when the user walks over narrow or curved paths. Data fusion techniques combining GPS and dead-reckoning could possibly improve accuracy in those nonideal conditions. Further data analysis could be performed to understand if oxygen saturation measurements from different pulse oximeters converge when patients are not moving. It is likely that the measurements are more reliable at least before the start and after the end of the test. If pulse oximetry during the test is desirable, motion-resistant pulse oximeters may provide a reliable solution for this scenario [27]. Compliance to the testing regime could be boosted by simple techniques like reminders. The 6MWD could be linked or complemented with data collected by wearables and fitness trackers, which may provide a less obtrusive way of measuring exercise capacity.

In terms of clinical significance, longitudinal studies in which the data collected by the app are used in clinical practice are needed to understand the usefulness of these tests. For example, could remote monitoring improve follow-up and support titration of medications? We also expect that clinical use of the app

would affect compliance. How would patients behave when they know their use of the app could lead to different health outcomes?

Limitations

This pilot study involved 30 patients and was not aimed at obtaining statistical significance; larger cohorts would be needed to confirm the results. It is also important to observe that our manual data quality assurance strategy was not always consistent and that an unestimated number of outdoor tests performed in the wrong conditions were included in the statistics.

Conclusions

Our app-based outdoor 6MWT in community settings is valid, repeatable, and well accepted by patients. Its use could complement or potentially substitute conventional 6MWTs in clinics. The same app, however, is not accurate enough for clinical use when used indoors.

Acknowledgments

The research described in this paper was supported by the National Institute for Health Research Biomedical Research Centre, Oxford. The data that support the findings of the study are available upon reasonable request from the corresponding author. The data is not publicly available due to containment of personal health information. The algorithms used in the app have been made publicly available as an appendix in Salvi et al [14]. A simplified version of the app that uses the same outdoor distance estimation algorithm has been released by one of the authors [28]. Code used for data analysis can be made available to reasonable requests.

Authors' Contributions

Contributions are listed according to the CRediT taxonomy. DS, LT, and EO contributed to the conceptualization and methodology of the project. DS developed the software, performed data curation and analysis, and managed the project. EP helped with data acquisition and the investigation. LT and EO supervised the project and helped acquire funding. All authors contributed to writing and reviewing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of valid and invalid tests.
[DOCX File , 131 KB - mhealth_v9i6e22748_app1.docx]

Multimedia Appendix 2
Usability and acceptance questionnaire.
[DOCX File , 15 KB - mhealth_v9i6e22748_app2.docx]

Multimedia Appendix 3
Interview questions.
[DOCX File , 13 KB - mhealth_v9i6e22748_app3.docx]

References

1. Agarwal R, Gomberg-Maitland M. Current therapeutics and practical management strategies for pulmonary arterial hypertension. Am Heart J 2011 Aug;162(2):201-213. [doi: 10.1016/j.ahj.2011.05.012] [Medline: 21835279]
2. Galiè N, Brundage BH, Ghofrani HA, Oudiz RJ, Simonneau G, Safdar Z, Pulmonary Arterial HypertensionResponse to Tadalafil (PHIRST) Study Group. Tadalafil therapy for pulmonary arterial hypertension. Circulation 2009 Jun 09;119(22):2894-2903. [doi: 10.1161/CIRCULATIONAHA.108.839274] [Medline: 19470885]
3. Peacock AJ, Murphy NF, McMurray JJV, Caballero L, Stewart S. An epidemiological study of pulmonary arterial hypertension. Eur Respir J 2007 Jul;30(1):104-109 [FREE Full text] [doi: 10.1183/09031936.00092306] [Medline: 17360728]
4. Enright PL. The six-minute walk test. Respir Care 2003 Aug;48(8):783-785 [FREE Full text] [Medline: 12890299]
5. Gabler NB, French B, Strom BL, Palevsky HI, Taichman DB, Kawut SM, et al. Validation of 6-minute walk distance as a surrogate end point in pulmonary arterial hypertension trials. Circulation 2012 Jul 17;126(3):349-356 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.105890] [Medline: 22696079]
6. Benza RL, Miller DP, Gomberg-Maitland M, Frantz RP, Foreman AJ, Coffey CS, et al. Predicting survival in pulmonary arterial hypertension. Circulation 2010 Jul 13;122(2):164-172. [doi: 10.1161/circulationaha.109.898122]
7. Macchia A, Marchioli R, Marfisi R, Scarano M, Levantesi G, Tavazzi L, et al. A meta-analysis of trials of pulmonary hypertension: a clinical condition looking for drugs and research methodology. Am Heart J 2007 Jun;153(6):1037-1047. [doi: 10.1016/j.ahj.2007.02.037] [Medline: 17540207]

8. Miyamoto S, Nagaya N, Satoh T, Kyotani S, Sakamaki F, Fujita M, et al. Clinical correlates and prognostic significance of six-minute walk test in patients with primary pulmonary hypertension. Comparison with cardiopulmonary exercise testing. Am J Respir Crit Care Med 2000 Feb;161(2 Pt 1):487-492. [doi: 10.1164/ajrccm.161.2.9906015] [Medline: 10673190]
9. Barst RJ, Rubin LJ, Long WA, McGoon MD, Rich S, Badesch DB, Primary Pulmonary Hypertension Study Group. A comparison of continuous intravenous epoprostenol (prostacyclin) with conventional therapy for primary pulmonary hypertension. N Engl J Med 1996 Feb 01;334(5):296-301. [doi: 10.1056/NEJM199602013340504] [Medline: 8532025]
10. van de Port I, Wevers L, Kwakkel G. Is outdoor use of the six-minute walk test with a global positioning system in stroke patients' own neighbourhoods reproducible and valid? J Rehabil Med 2011 Nov;43(11):1027-1031. [doi: 10.2340/16501977-0881]
11. Jehn M, Schmidt-Trucksäess A, Schuster T, Hanssen H, Weis M, Halle M, et al. Accelerometer-based quantification of 6-minute walk test performance in patients with chronic heart failure: applicability in telemedicine. J Card Fail 2009 May;15(4):334-340. [doi: 10.1016/j.cardfail.2008.11.011] [Medline: 19398082]
12. Ata R, Gandhi N, Rasmussen H, El-Gabalawy O, Gutierrez S, Ahmad A, et al. Clinical validation of smartphone-based activity tracking in peripheral artery disease patients. NPJ Digit Med 2018 Dec 11;1(1):1. [doi: 10.1038/s41746-018-0073-x]
13. 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] [Medline: 26283292]
14. Salvi D, Poffley E, Orchard E, Tarassenko L. The mobile-based 6-minute walk test: usability study and algorithm development and validation. JMIR Mhealth Uhealth 2020 Jan 03;8(1):e13756 [FREE Full text] [doi: 10.2196/13756] [Medline: 31899457]
15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
16. Denegar C, Ball D. Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement. J Sport Rehab 1993;2(1):35-42. [doi: 10.1123/jsr.2.1.35]
17. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
18. Sun Y, Wang N, Peng Z. Understanding the acceptance of mobile health services: a comparison and integration of alternative models. J Electr Commerce Res 2013;14(2):183-200 [FREE Full text]
19. Simonneau G, Montani D, Celermajer D, Denton C, Gatzoulis M, Krowka M, et al. Haemodynamic definitions and updated clinical classification of pulmonary hypertension. Eur Respir J 2019 Jan;53(1):1 [FREE Full text] [doi: 10.1183/13993003.01913-2018] [Medline: 30545968]
20. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What predicts patients' adoption intention toward mHealth services in China: empirical study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]
21. Mainguy V, Malenfant S, Neyron A, Bonnet S, Maltais F, Saey D, et al. Repeatability and responsiveness of exercise tests in pulmonary arterial hypertension. Eur Respir J 2013 Aug;42(2):425-434 [FREE Full text] [doi: 10.1183/09031936.00107012] [Medline: 23100508]
22. Maeder A, Poultney N, Morgan G, Lippiatt R. Patient compliance in home-based self-care telehealth projects. J Telemed Telecare 2015 Dec;21(8):439-442. [doi: 10.1177/1357633X15612382] [Medline: 26556057]
23. Scalise L, Marchionni P, Carnielli V. Rejection of false saturation data in optical pulse-oximeter. Proc SPIE Int Soc Optics Photonics 2010;7715:1. [doi: 10.1117/12.854474]
24. Naraharisetti K, Bawa M. Comparison of different signal processing methods for reducing artifacts from photoplethysmograph signal. 2011 Presented at: IEEE International Conference on Electro/Information Technology; 2011; Mankato p. 1-8. [doi: 10.1109/eit.2011.5978571]
25. Lee J, Jung W, Kang I, Kim Y, Lee G. Design of filter to reject motion artifact of pulse oximetry. Comput Stand Interfaces 2004 May;26(3):241-249. [doi: 10.1016/s0920-5489(03)00077-1]
26. Randell C, Djiallis C, Muller H. Personal position measurement using dead reckoning. Seventh IEEE International Symposium on Wearable Computers 2003:166. [doi: 10.1109/iswc.2003.1241408]
27. Barker SJ. "Motion-resistant" pulse oximetry: a comparison of new and old models. Anesth Analg 2002 Oct;95(4):967-972. [doi: 10.1097/00000539-200210000-00033] [Medline: 12351278]
28. Simplified version of SMWT app. URL: https://github.com/dariosalvi78/TimedWalkApp [accessed 2021-05-23]

Abbreviations

6MWD: 6-minute walk test distance
6MWT: 6-minute walk test
HR: heart rate
PAH: pulmonary arterial hypertension

SpO_2: peripheral oxygen saturation

Edited by L Buis; submitted 22.07.20; peer-reviewed by J Cleland, G Darville; comments to author 15.09.20; revised version received 13.03.21; accepted 08.04.21; published 07.06.21.

Please cite as:
Salvi D, Poffley E, Tarassenko L, Orchard E
App-Based Versus Standard Six-Minute Walk Test in Pulmonary Hypertension: Mixed Methods Study
JMIR Mhealth Uhealth 2021;9(6):e22748
URL: https://mhealth.jmir.org/2021/6/e22748
doi:10.2196/22748
PMID:34096876

Original Paper

Development of an Android-Based Self-Report Assessment for Elderly Driving Risk (SAFE-DR) App: Mixed Methods Study

Ho Sung Hwang[1], MPH; Seong-Youl Choi[2], PhD

[1]Department of Occupational Therapy, Wonkwang University Gwangju Medical Center, Gwangju, Republic of Korea

[2]Department of Occupational Therapy, Gwangju Women's University, Gwangju, Republic of Korea

Corresponding Author:
Seong-Youl Choi, PhD
Department of Occupational Therapy
Gwangju Women's University
201, Yeodae-gil, Gwangsan-gu
Gwangju,
Republic of Korea
Phone: 82 10 3209 ext 9146
Email: ckshrj6@hanmail.net

Abstract

Background: Self-report assessments for elderly drivers are used in various countries for accessible, widespread self-monitoring of driving ability in the elderly population. Likewise, in South Korea, a paper-based Self-Report Assessment for Elderly Driving Risk (SAFE-DR) has been developed. Here, we implemented the SAFE-DR through an Android app, which provides the advantages of accessibility, convenience, and provision of diverse information, and verified its reliability and validity.

Objective: This study tested the validity and reliability of a mobile app-based version of a self-report assessment for elderly persons contextualized to the South Korean culture and compared it with a paper-based test.

Methods: In this mixed methods study, we recruited and interviewed 567 elderly drivers (aged 65 years and older) between August 2018 and May 2019. For participants who provided consent, the app-based test was repeated after 2 weeks and an additional paper-based test (Driver 65 Plus test) was administered. Using the collected data, we analyzed the reliability and validity of the app-based SAFE-DR. The internal consistency of provisional items in each subdomain of the SAFE-DR and the test-retest stability were analyzed to examine reliability. Exploratory factor analysis was performed to examine the validity of the subdomain configuration. To verify the appropriateness of using an app-based test for older drivers possibly unfamiliar with mobile technology, the correlation between the results of the SAFE-DR app and the paper-based offline test was also analyzed.

Results: In the reliability analysis, Cronbach α for all items was 0.975 and the correlation of each item with the overall score ranged from r=0.520 to r=0.823; 4 items with low correlations were removed from each of the subdomains. In the retest after 2 weeks, the mean correlation coefficient across all items was r=0.951, showing very high reliability. Exploratory factor analysis on 40 of the 44 items established 5 subdomains: on-road (8 items), coping (16 items), cognitive functions (5 items), general conditions (8 items), and medical health (3 items). A very strong negative correlation of –0.864 was observed between the total score for the app-based SAFE-DR and the paper-based Driver 65 Plus with decorrelation scales. The app-based test was found to be reliable.

Conclusions: In this study, we developed an app-based self-report assessment tool for elderly drivers and tested its reliability and validity. This app can help elderly individuals easily assess their own driving skills. Therefore, this assessment can be used to educate drivers and for preventive screening for elderly drivers who want to renew their driver's licenses in South Korea. In addition, the app can contribute to safe driving among elderly drivers.

(JMIR Mhealth Uhealth 2021;9(6):e25310) doi:10.2196/25310

KEYWORDS

Android driving app; driving safety; reliability; self-assessment; validity; mHealth; driving

Introduction

Driving increases the autonomy of elderly individuals, which helps them to participate in activities in the local community and at home [1]. Moreover, elderly drivers who have higher independence and community participation show higher life satisfaction [2]; however, aging is accompanied by factors that impede safe driving, such as perceptual, motor, and cognitive impairments [3,4]. Due to health decline, elderly drivers have much higher accident rates than all other age groups and are more likely to suffer severe injury and death in the event of an accident [5,6]. Therefore, it is becoming increasingly important to develop ways of testing their health status and assessing their suitability for driving [7].

Elderly drivers have the self-control to reflect on their driving ability and prevent driving risk themselves [8]. They generally tend to avoid dangerous driving situations, such as long-duration driving, high-speed driving, and uncomfortable driving situations such as driving at night or in the rain [9]. Using these characteristics of elderly drivers, continuing research is being conducted to develop various self-report assessments and test their correlation with driving risks [10].

Self-report assessment plays an important role in elderly drivers' decision-making on the continuation of driving and also presents the following advantages: individuals can test their status using a relatively simple method, the assessment can be distributed to a broad population online, and it provides rapid feedback [11]. Existing self-report assessments for elderly drivers were previously only available as paper-based tests; they are now available on CD-ROM [12] and via websites [13]. As such, self-assessment services for elderly drivers are shifting away from the conventional method of paper-based testing to a more efficient way that fits social reality, and the new service modalities should be able to provide older drivers with easier access and faster feedback.

Recently, numerous mobile-based health care apps have been developed to aid health care across a wide population [14]. Although the elderly population may have restricted access to mobile-based services, this is outweighed by various benefits, such as diverse access to information, convenience, and reduced social isolation [15]. Mobile-based services for self-report assessments for elderly drivers have not been attempted so far because elderly persons are not familiar with the mobile environment. Nevertheless, according to the mobile technology acceptance model for the elderly aged 65 years and older, they are strongly influenced by surrounding reference social groups (children and relatives) and do not show major differences from other age groups in motivation to access mobile devices [16]. In other words, even though elderly drivers may experience some difficulties performing mobile-based self-report assessments, it is important to investigate the usability of new methods that provide easier access and feedback.

In the United States, Austria, and most European countries, screening policies based on driving self-assessments are widely used for driver's license renewals for the elderly population; assessments include the Driving Decisions Workbook, Older Driver's Self-Assessment Questionnaire, Driving Safely While Aging Gracefully, Safe Driving Behavior Measure, and Driver 65 Plus, among others [17-20]. South Korea also has a traffic safety policy that shortens the length of license renewal and provides professional evaluations for elderly drivers [21]. Within these policies, a means of rapidly screening the driving risk for a wide range of elderly people is needed.

Patterns of driving behavior vary depending on sociocultural characteristics, and elderly drivers' driving abilities are strongly influenced by perceptual ability and culture [22]. Therefore, it is necessary to thoroughly examine the reliability and validity of self-report assessments in which elderly drivers examine their own status within the corresponding sociocultural environment. To this end, self-report assessments have been developed to suit the characteristics of various countries, such as the United States, Canada, and Australia [11,23]. Likewise, the Self-Report Assessment For Elderly Driving Risk (SAFE-DR) was developed as a self-report assessment for elderly drivers suited to the cultural characteristics of South Korea (Multimedia Appendix 1) [24].

With the above evidence, the purpose of this study was to develop and test a mobile version of a self-report assessment app for elderly persons contextualized to the South Korean culture. We then tested its validity and reliability by comparing its results to that of an established, validated, paper-based test.

Methods

Study Design

This was a mixed methods study that included a qualitative exploratory method of a focus group to develop a mobile-based version of the SAFE-DR and a cross-sectional survey conducted among elderly South Korean drivers to assess the validity and reliability of a self-administered Android-based mobile app in assessing their driving risk. Before starting the study, the entire protocol was approved by the Kwangju Women's University Institutional Review Board (No. 1041485-201709-HR-001-29). All participants received and understood an oral and written explanation of the entire study and provided written informed consent to participate.

Procedure

This study constituted three phases: (1) development and implementation of the SAFE-DR app, (2) testing of the reliability and validity of the SAFE-DR app, and (3) analysis of the correlation between the SAFE-DR app and a paper-based self-report assessment for elderly drivers.

Phase One: Implementing the SAFE-DR App

App Configuration

To implement the SAFE-DR as an Android app, we used the qualitative exploratory method of focus group interactions [25]. The focus group consisted of 2 app developers and 3 professors specializing in elderly driving rehabilitation. The investigators in this study directed communications within the focus group; the professors specializing in elderly driving rehabilitation composed app scenarios, and the 2 app developers used the Android Studio to produce a pilot file based on these scenarios. App scenarios consisted of using functional and timing point

of views to describe or visualize interactions between users and mobile systems, and app developers anticipated use episodes before coding into computer languages. Elderly driving experts discussed scenario visualization, content application on a page, score expression as an evaluative result, and explanations to users about items with low scores. The scenario, including the contents discussed and the message sequence chart, was written using the Excel 2016 program (Microsoft Corp) for each app page. Based on this scenario, the app development specialists used a programming language (Java) to construct the app. Next, the whole group participated in a process of qualitative communication to finalize the function and design of the app from the users' perspective. The app was designed in 4 sections: explanation of evaluation, guidance, and consent to use; data entry of personal and driving characteristics; performance of test through items of each subdomain; and presentation of test results. The test can be completed in about 20 minutes; the app analyzes the subject's response to each item (based on a 3-point scale) to provide information on the user's driving ability. It further cautions the driver to be careful or see an expert for a driving ability test. Finally, an app administration website was created for data management.

SAFE-DR

The SAFE-DR was developed by SYC to reflect the characteristics of elderly drivers in South Korea [24]. Via a Delphi survey, a 28-person panel selected 44 assessment items and, after analysis of offline reliability and validity, 38 items were included in the final instrument, with subdomains of on-road, coping, and health. Each item was measured on a 3-point Likert scale of agree, disagree, and strongly disagree. The reliability coefficients for all items (on-road, coping, and health subdomains) were 0.906, 0.921, and 0.913, respectively. The instrument demonstrated content, construct, and predictive validity. In the implementation of the app version of the test in this study, we used the 44 items that showed content validity in the paper-based SAFE-DR, collected data from elderly drivers, and assessed the test items. The app provides information to elderly drivers on any reduction in driving abilities and whether they should discuss their abilities with a driving expert. Based on this information, elderly drivers will be able to make decisions as to whether they should seek expert evaluation or exercise caution in situations to practice safe driving.

Phase Two: Testing the Reliability and Validity of the SAFE-DR App

Testing the Reliability and Validity

Construct validity was analyzed to ensure that the hypotheses for the subdomain configuration of SAFE-DR developed offline were consistent with the actual app-based assessment. The reliability of the configuration items for each subarea was analyzed to select items that were inconsistent. To analyze test-retest reliability, the SAFE-DR app was readministered 2 weeks later to the 81 participants who consented to the retest.

Participants and Data Collection

Data were collected between August 2018 and May 2019 through the SAFE-DR app made by the focus group and through one-to-one interviews with elderly drivers. Table 1 shows the general characteristics of the focus group. The focus group developed the app by discussing the topics of organizing app scenarios (function, point of view, interaction, and visualization) and organizing elderly driver assessment content (expression of evaluation content, items, and results).

Table 1. General characteristics of the focus group (n=5).

Characteristic	Value
Sex, n (%)	
Male	4 (80)
Female	1 (20)
Age (years), mean (SD)	42.60 (9.63)
35-40, n (%)	3 (60)
40-45, n (%)	1 (20)
45-50, n (%)	0 (0)
50-55, n (%)	0 (0)
55-59, n (%)	1 (20)
Area of expertise, n (%)	
App development	2 (40)
Elderly driving rehabilitation	3 (60)
Experience with app development or elderly driving rehabilitation (years), n (%)	
5-10	1 (20)
10-15	3 (60)
15-20	0 (0)
20-25	1 (20)

Since it is difficult to specify the population of elderly drivers in South Korea, we used convenience sampling and snowball sampling. After obtaining consent from 3 public health centers and 3 general hospitals in Seoul, Gyeonggi-do, and Gwangju, which have large elderly dynamic populations, we installed testing booths and advertised the study to recruit elderly drivers. In addition, we recruited further participants by advertising by word of mouth through the elderly drivers who had taken the test. All individuals who participated in the study received written and verbal guidance for the entire research process and provided written informed consent to participate in the study.

Participants were enrolled between August 2018 and May 2019, and a total of 567 elderly drivers participated in the study. All participants were interviewed to check general and driving-related characteristics; they completed the SAFE-DR assessment themselves using the developed Android app. In addition, participants who consented to further testing completed an offline self-report assessment using the Driver 65 Plus. Individuals met the inclusion criteria if they were individuals aged 65 years and over who possessed a driving license and had at least 1 year of driving experience and whose cognitive and verbal abilities imposed no restrictions on completing the app-based self-report assessment. Based on the inclusion criteria, of the 567 participants, 51 persons were excluded because they were under 65 years of age or provided incomplete responses to the self-report assessment. The remaining 516 participants were included in the final analysis.

Phase Three: Analyzing the Correlation Between the SAFE-DR App and a Paper-Based Self-Report Assessment for Elderly Drivers (Driver 65 Plus Test)

Correlation Between App and Paper-Based Assessment

The third phase was to test the validity of the assessment results of the elderly drivers, who may not be familiar with mobile apps, by analyzing the correlation between the results of the SAFE-DR app and a paper-based offline test. Among the elderly drivers who used the SAFE-DR app, those who consented to further testing were asked to complete the Korean version of the Driver 65 Plus test, which is one of the most widely used offline self-report assessments for elderly drivers. The correlation between the SAFE-DR app and paper-based Driver 65 Plus results was analyzed to verify consistency between the results of the offline and online tests. In the Driver 65 Plus test, lower scores indicated a more positive result, whereas in the SAFE-DR test, higher scores indicated a more positive result. Therefore, we expected to observe a negative value for the correlation coefficient between the two tests.

Driver 65 Plus

This test was developed by the American Automobile Association based on the Safety Research and Education Project of the Teacher's College of Columbia University [26]. A feature of this test is that the scoring scale for each item is presented differently depending on the content such that elderly drivers cannot predict the question's intentions. The Korean Driver 65 Plus has shown test-retest reliability of .95 and construct and concurrent validity [27].

Statistical Analysis

For data analysis, we used PASW Statistics (version 18.0, IBM Corp). Participant characteristics were analyzed using the descriptive statistics of mean, frequency, and percentage. Cronbach α and the correlation among all items were analyzed to examine the internal consistency of provisional items in each subdomain of the SAFE-DR and the stability of a retest after 2 weeks.

Exploratory factor analysis was performed to classify the subdomains of SAFE-DR. For exploratory factor analysis, we used principal component analysis, selecting factors with an initial eigenvalue ≥1.0, and used a varimax rotation. To interpret the results, we checked whether the Kaiser-Meyer-Olkin value in Bartlett test of sphericity was ≥0.8, that the factor loading of each item was ≥0.4, and that the eigenvalue of each factor was ≥1.0. Although a factor loading ≥0.5 for each item is ideal, given that this was a sociological study, we interpreted the results using a criterion of 0.4 [28,29]. Pearson correlation analysis was used to estimate the correlation coefficients for each item in test-retest reliability and to estimate the correlation between the paper-based self-report assessment results and the app-based self-report assessment results.

Results

General Characteristics of Participants

Table 2 shows the general characteristics of the elderly drivers who participated in this study. Of the elderly drivers, 79.7% (411/516) were male and 20.3% (105/516) were female, while the mean age was 72.03 years. A total of 45% (232/516) were taking regular medication for hypertension (187/516, 80.6%), diabetes (25/516, 10.8%), hyperlipidemia (11/516, 4.7%), and prostate disorders (6/516, 2.6%); however, 3 drivers (3/516, 1.3%) were also taking antiepileptic medication, which could affect driving. The automobile types most commonly used by the elderly drivers were sport utility vehicles (168/516, 32.6%), midsize sedans (144/516, 27.9%), and compact cars (114/516, 22.1%). The mean driving time of the 516 participants was 22.75 years, the average driving time per week for the 479 currently driving drivers was 4.49 hours, and the average driving cessation period for the 37 participants who were not currently driving was 1.78 years. A total of 4.4% (21/516) had experienced driving accidents during the last 3 months, of which the majority (13/516, 2.5%) experienced a collision due to their own fault.

Table 2. General characteristics of the elderly drivers (n=516).

Characteristic	Value
Sex, n (%)	
Male	411 (79.7)
Female	105 (20.3)
Age (years), mean (SD)	72.03 (5.13)
65-69, n (%)	192 (37.2)
70-79, n (%)	269 (52.1)
≥80, n (%)	55 (10.7)
Regular medicine, n (%)	
Yes	232 (45.0)
Hypertension	187 (80.5)
Diabetes	25 (10.8)
Hyperlipidemia	11 (4.7)
Prostate problems	6 (2.6)
Antiepileptic	3 (1.3)
No	284 (55.0)
Automobile type, n (%)	
Compact car	114 (22.1)
Midsize sedan	144 (27.9)
Sport utility vehicle	168 (32.6)
Large size sedan	35 (6.8)
Business vehicle	3 (0.6)
Other	52 (10.1)
Driving experience (years), mean (SD)	22.75 (8.04)
Currently driving	
Yes, n (%)	479 (92.8)
Total driving time per week (hours), mean (SD)	4.37 (2.36)
No, n (%)	37 (7.2)
Period (years), mean (SD)	1.78 (1.45)
Driving accident in the past 3 months, n (%)	
Yes	21 (4.1)
Individual accident	1 (0.2)
Minor accident (by me)	3 (0.6)
Minor accident (by others)	2 (0.4)
Collisions (by me)	13 (2.5)
Collisions (by others)	2 (0.4)
No	495 (95.9)

Internal Consistency

To analyze the internal consistency of the SAFE-DR app, we analyzed Cronbach α and the correlation of each item with the overall score. For all items, Cronbach α=.975, and the correlation of each item with the overall score ranged from r=.520 to r=.823, meaning that all items showed a correlation of at least .50 (Table 3).

Table 3. Correlation of individual items to the total Self-Report Assessment for Elderly Driving Risk score and test-retest reliability (n=516).

Pre-subarea and item number	Mean (SD)	Item to total correlation	Cronbach α		Total[a]
			When item is deleted	Test-retest[b] (n=101)	
On-road	—[c]	—	—	—	.921
1	2.44 (.664)	.418	.924	.932	—
2	2.35 (.721)	.556	.919	.965	—
3	2.35 (.736)	.709	.913	.963	—
4	2.09 (.806)	.738	.912	.961	—
5	2.10 (.823)	.771	.910	.965	—
6	2.19 (.781)	.804	.909	.933	—
7	2.40 (.688)	.771	.911	.941	—
8	2.41 (.672)	.801	.910	.944	—
9	2.27 (.764)	.642	.916	.947	—
10	2.27 (.780)	.703	.913	.925	—
11	2.30 (.729)	.756	.911	.947	—
12	2.20 (.739)	.413	.925	.936	—
Coping	—	—	—	—	.946
13	2.34 (.679)	.509	.947	.976	—
14	2.48 (.628)	.614	.945	.951	—
15	2.26 (.732)	.687	.943	.930	—
16	2.30 (.732)	.739	.942	.960	—
17	2.25 (.722)	.762	.942	.940	—
18	2.38 (.678)	.551	.946	.919	—
19	2.21 (.762)	.646	.944	.980	—
20	2.26 (.726)	.756	.942	.966	—
21	2.34 (.672)	.772	.942	.918	—
22	2.08 (.793)	.685	.944	.936	—
23	2.25 (.742)	.799	.941	.961	—
24	2.38 (.667)	.782	.942	.945	—
25	2.19 (.784)	.733	.942	.954	—
26	2.23 (.751)	.792	.941	.957	—
27	2.21 (.783)	.722	.943	.952	—
28	2.25 (.758)	.714	.943	.980	—
Health	—	—	—	—	.936
29	2.23 (.775)	.629	.934	.979	—
30	2.20 (.753)	.714	.931	.970	—
31	2.23 (.726)	.680	.932	.941	—
32	2.29 (.718)	.712	.931	.982	—
33	2.41 (.658)	.715	.931	.974	—
34	2.04 (.777)	.629	.934	.942	—
35	2.23 (.730)	.666	.933	.981	—
36	2.14 (.769)	.691	.932	.948	—
37	1.99 (.809)	.682	.932	.945	—

Pre-subarea and item number	Mean (SD)	Item to total correlation	Cronbach α		Total[a]
			When item is deleted	Test-retest[b] (n=101)	
38	2.03 (.795)	.721	.931	.976	—
39	2.24 (.754)	.750	.930	.952	—
40	2.22 (.741)	.708	.931	.938	—
41	2.28 (.739)	.738	.931	.964	—
42	2.59 (.585)	.551	.935	.974	—
43	2.54 (.633)	.604	.934	.909	—
44	2.67 (.520)	.509	.936	.898	—

[a]Total Cronbach α=.975.

[b]Mean of the total test-retest correlations, r=.951.

[c]Not applicable.

For each of the subdomains defined when the SAFE-DR app was constructed, any item that increased Cronbach α when removed was excluded from the assessment; 5 items showed either increased or equal reliability when removed. For items that showed the same reliability, we made decisions regarding deletion based on subdomain composition in the factor analysis. As a result of this process, we decided to remove items 1, 12, 13, and 18.

Construct Validity

In the exploratory factor analysis, the Kaiser-Meyer-Olkin value was 0.963, which was ≥0.8 and close to 1, and the eigenvalues of each factor were ≥1.0, which was independently sufficient to form subdomains. The result of Bartlett test of sphericity, which indicates the suitability of the total data in the factor model, was statistically significant (P<.001, chi-square: 17,265.9; Table 4).

Table 4. Exploratory factor analysis for Self-Report Assessment for Elderly Driving Risk subdomains.

Final subarea and item number	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Communality summary loading
On-road						
2	–.028	.059	.692	.275	.260	.626
3	.200	.153	.719	.240	.229	.690
4	.420	.587	.438	.096	.040	.724
5	.377	.477	.574	.131	.066	.721
6	.392	.378	.652	.149	.054	.747
7	.384	.170	.639	.306	.211	.722
8	.379	.241	.641	.273	.187	.722
9	.331	.426	.510	–.033	.177	.584
Coping						
10	.560	.237	.426	.230	.145	.624
11	.542	.346	.393	.190	.267	.675
14	.405	.053	.330	.187	.514	.575
15	.536	.222	.396	.146	.158	.540
16	.537	.239	.513	.158	.211	.678
17	.513	.422	.351	.247	.176	.656
19	.527	.388	.105	.082	.250	.509
20	.656	.283	.161	.289	.177	.651
21	.564	.294	.333	.280	.227	.645
22	.465	.604	.158	.125	.083	.629
23	.600	.399	.253	.307	.163	.704
24	.540	.329	.301	.366	.208	.668
25	.662	.307	.141	.283	.172	.662
26	.649	.443	.223	.190	.157	.728
27	.682	.370	.159	.175	.046	.660
28	.531	.271	.342	.401	.090	.642
Health						
Cognitive functions						
29	.157	.126	.366	.737	.156	.742
30	.447	.275	.337	.505	.122	.659
31	.265	.251	.139	.731	.166	.714
32	.231	.280	.251	.689	.199	.709
33	.311	.242	.140	.602	.378	.680
General conditions						
34	.220	.659	.165	.141	.193	.567
35	.202	.596	.204	.287	.217	.567
36	.233	.703	.205	.200	.166	.658
37	.212	.799	.135	.181	.098	.744
38	.369	.681	.162	.254	.084	.698
39	.372	.516	.138	.478	.118	.666
40	.358	.476	.221	.370	.165	.569
41	.428	.551	.157	.231	.290	.649

Final subarea and item number	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Communality summary loading
Medical health						
42	.161	.175	.186	.158	.784	.730
43	.158	.284	.174	.187	.735	.712
44	.123	.104	.177	.160	.843	.795
Eigenvalue	20.585	2.124	1.574	1.324	1.033	—[a]
Variance explained (%)	51.462	5.310	3.935	3.309	2.581	—
Cumulative variance (%)	51.462	56.773	60.707	64.017	66.598	—

[a]Not applicable.

Factor Exploration

Although a factor loading ≥0.5 for each item is ideal, given that this was a sociological study, we derived a solution that satisfies factor loading ≥0.4. Considering the semantic units of individual items, we decided on a factor solution that did not disturb the semantic units of items included in the theoretically chosen subdomains of on-road, coping, and health. Therefore, during factor analysis, the input arrangement of items was used instead of ordering the items by factor component size. As a result, in the final factor matrix, the initial 3 putative subdomains were divided into 5 subdomains. The on-road and coping subdomains maintained their theoretical arrangements, whereas the items included in the theoretical health subdomain were divided into 5 subdomains. Reviewing the semantic content of the grouped items, we defined items 29-33 as the cognitive functions subdomain, items 34-41 as the general conditions subdomain, and items 42-44 as the medical health subdomain (Table 3).

Stability

To analyze measurement stability using the SAFE-DR app, we retested the app using the same test after 2 weeks. We analyzed the correlation between initial test results and retest results, and the correlation coefficient ranged from r=.898 to r=.982. The mean correlation coefficient across all items was r=.951 (Table 2).

Correlation Between App-based SAFE-DR and Driver 65 Plus

To investigate whether elderly drivers, who may not be accustomed to a mobile environment, showed similar trends in the results of the SAFE-DR app and that of the paper-based test, we analyzed the correlation between Driver 65 Plus scores and the total SAFE-DR score with the subdomains. The total SAFE-DR score showed a negative correlation of –0.864 with the Driver 65 Plus score; among subdomains, coping showed the strongest correlation (–0.812). A Pearson correlation coefficient ≥0.8 indicates a very strong correlation. Meanwhile, the on-road, cognitive functions, and general conditions subdomains also showed strong correlations of –0.768, –0.758, and –0.767, respectively. The medical health subdomain showed a significant but moderate correlation of –0.456 (Table 5). Therefore, elderly drivers did not show especially different responses in the app-based method from the paper-based method.

Table 5. Correlation between Self-Report Assessment for Elderly Driving Risk and Driver 65 Plus (n=81).

Topic	Driver 65 Plus	App-based SAFE-DR[a]					
		On-road	Coping	Cognitive functions	General conditions	Medical health	Total
Driver 65 Plus	1	—[b]	—	—	—	—	—
On-road	–0.768	1	—	—	—	—	—
Coping	–0.812	0.807	1	—	—	—	—
Cognitive functions	–0.758	0.630	0.788	1	—	—	—
General conditions	–0.767	0.732	0.786	0.758	1	—	—
Medical health	–0.456	0.381	0.414	0.339	0.343	1	—
Total	–0.864	0.882	0.963	0.845	0.887	0.479	1

[a]Self-Report Assessment for Elderly Driving Risk.

[b]Not applicable.

Final App for SAFE-DR

The final app for SAFE-DR is easily accessible through the Google Play app store. Examples of app screens are shown in Figures 1 and 2.

Figure 1. Title and information collection screenshot of the Android app: title page (left) and driver-related information such as gender, date of birth, residence, license, driving status, primary means of transportation, and accident during the past 3 months (right).

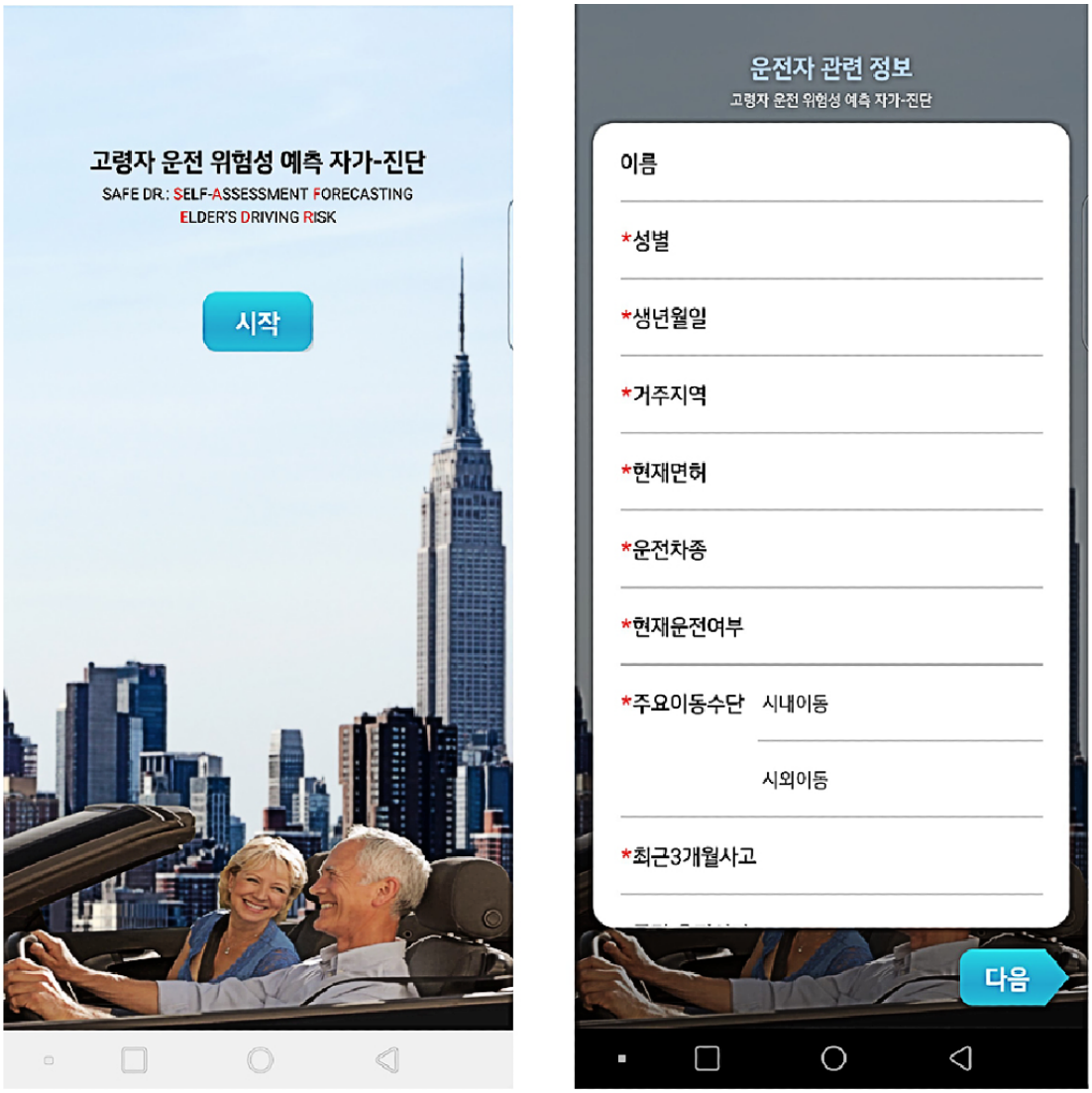
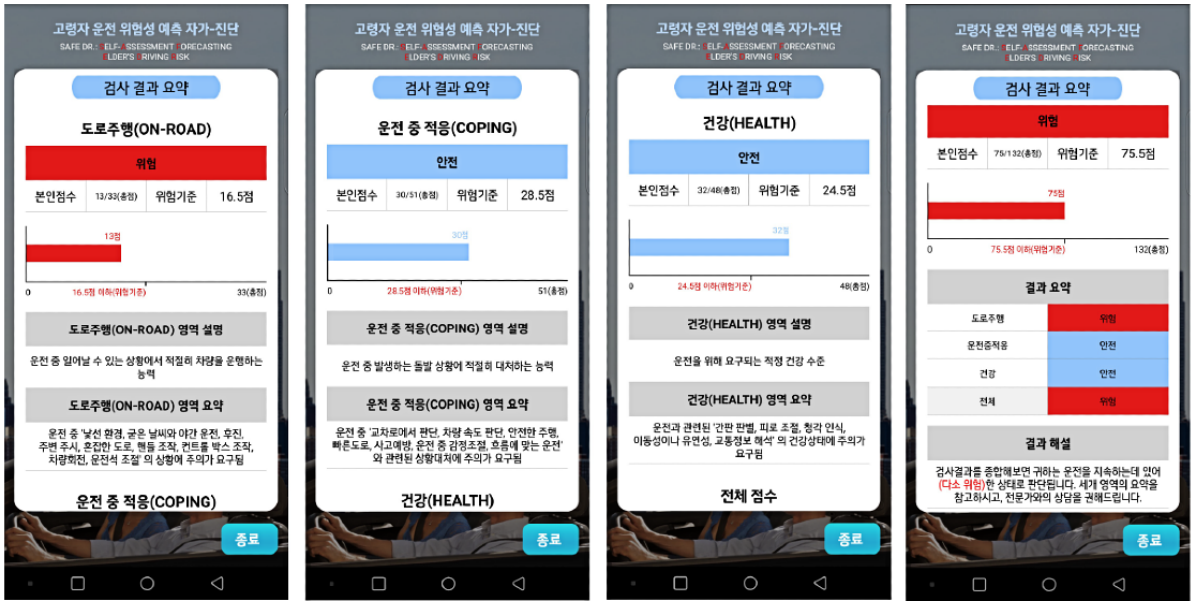


Figure 2. Screenshot of test results in the Android app. Displayed are subarea, total test results (classification of risk or safety in driving, graph of scores against reference points), and information requiring attention.



Discussion

Principal Findings

In this study, we developed the SAFE-DR as an app-based self-report assessment for elderly drivers in South Korea. We tested the reliability and validity of the SAFE-DR in a mobile environment and demonstrated that the app-based testing method can be effectively used for elderly drivers.

For the development of the conventional paper-based SAFE-DR, 44 items were tested for content validity by an expert panel in a Delphi survey, and 38 items were selected in the subdomains of on-road, coping, and health. In the app-based SAFE-DR test in this study, 40 out of the 44 items were selected; items 9, 19, and 43 had been excluded from the paper-based test but were not excluded from the app-based test. In the factor analysis in our study, these items showed the lowest communality scores in their respective factors. In exploratory factor analysis, communality is a measure of how well each factor represents a given item [28]. Therefore, these 3 items can be interpreted as showing a relatively low correlation with other items in the same factor in the app-based test as well. Likewise, in the analysis of item reliability, these items tended to show lower reliability. When the paper-based test was developed, a sample of 339 participants was used, whereas, in our study, the sample size was 516 persons; reliability and validity analysis results are more likely to show higher values with a larger sample size [30]. Therefore, although the items showing low reliability and validity were similar between the paper-based and app-based tests, we can cautiously surmise that, due to the larger sample size, the results in our test had higher acceptance. The fact that the results of both the offline and online tests showed similar trends can be considered an important indicator in the development of the app-based test.

This study provides new findings as we implemented the SAFE-DR in an app, collected data, and verified its objectivity. Elderly drivers showed similar trends in the app-based test that were similar to using a paper-based, offline method. The results of the analysis showed a very strong negative correlation between the app-based SAFE-DR and the paper-based Driver 65 Plus; among subdomains of the SAFE-DR, medical health showed a moderate negative correlation and other subdomains showed strong negative correlations. These results suggest that the results of the app-based test have a linear relationship with the results of the paper-based test. Therefore, elderly drivers who may not be familiar with a mobile environment can use the app-based SAFE-DR without any major differences from the paper-based assessment. A study on the use of mobile-based services by elderly populations in various countries indicated that most elderly individuals are already prepared to use mobile-based services; their use reportedly improves independent living among the elderly [31]. Accordingly, it will be necessary to provide various services using a mobile environment such as apps for elderly drivers in Korea as well, including self-report assessments that can provide immediate and faster access to more information.

Some differences were observed between the paper-based test and our study's app-based test in subdomain composition. We first hypothesized theoretical factors based on the 3 subdomains of the basic paper-based SAFE-DR, and from the performed exploratory factor analysis, 5 factors emerged. This is because, among the putative subdomains, the health subdomain was split into 3 subdomains; these subdomains' items are related to fatigue during driving, driving-related cognition, visual perception, physical ability, and medical health status, including medication. The fact that health factors were further divided indicates that for the participants in our study the response to driving-related health status differed depending on specific health factors. Items 29-33, which were combined into one factor, were defined as the cognitive functions subdomain, as they shared content relating to memory, cognition, and geographical orientation. Items 34-41 were defined under the broad heading of general conditions, as they shared content on fatigue during driving, visual perception, auditory perception, physical fatigue, and flexibility. Finally, items 42-44 were defined as the medical health subdomain, as they showed content relating to medication, doctor consultations, epilepsy, and seizures.

The heterogeneity of digital literacy in the elderly population's use of mobile phones prevents older people from accessing smartphones. In a study on attitudes regarding mobile phone use in the daily lives of the elderly, there are individual differences in accessibility to mobile phones. If they refused to use a mobile phone, they felt uncomfortable reading the text on the touch screen. However, there were individuals who learned through personal networks and user manuals and used mobile phones [32]. In this respect, we tried to improve the visibility for older adults in app-based SAFE-DR configurations. Specifically, we increased the font size, reduced the number of characters displayed per screen, and encouraged the use of scroll functions. In a society where single-person mobile phone use has become routine, the elderly, like the populations of other age groups, can no longer avoid the use of smartphones. Therefore, the elderly should continue to try to use mobile devices. In this study, there was a difference between the paper-based evaluation and the app-based evaluation, but it was not a critical issue and it can be modified through feedback from elderly users. The app is currently being distributed through Google Play and is available free of charge. The researcher will continue to maintain this service for elderly drivers and plans to supplement it through feedback from elderly users.

Comparison With Prior Work

Regarding the subdomain compositions of previous self-report assessments for elderly drivers, the Driving Decisions Workbook developed in the United Kingdom contains on-road, seeing, thinking, getting around, and health subdomains [33]; the Self-Awareness and Feedback for Responsible Driving developed by the US Department of Transportation contains seeing, thinking, and getting around subdomains [13], and the Royal Automobile Club of Queensland Older Drivers' Self-Assessment Questionnaire developed by the Australian Automobile Association contains driving and health subdomains [34]. Among the subdomains in our study, on-road was included using the same name as previous assessments, and coping refers to the ability to cope with specific situations that can occur while driving on the road, which is consistent with the previous

getting around subdomain. For health-related items, the cognitive functions subdomain in our study has similar content to the items in the thinking subdomains in previous assessments. Conversely, the general conditions and medical health subdomains showed differences from those of previous assessments; items in the previous seeing subdomain were partially included in general conditions. Content in the medical health subdomain relating to medication or doctor consultations was previously included in the health subdomain in other assessments; however, since the division of roles between doctors and pharmacists in South Korea, the rate of patients seeking preventive medication guidance or consultations with doctors before disease or an accident has declined [35]. It is thought that these items were classified as a single factor, because unlike in other countries, elderly drivers showed different health status and different patterns in questions related to their medical state. This supports the idea that the app-based SAFE-DR test developed in this study properly reflects the cultural circumstances in South Korea.

The Australasian Model License Assessment Procedure states that the ideal procedure for driver's license renewal for elderly drivers follows a process of applying for testing through the local community, medical testing, multistage driving ability assessment, and provision of license options [36]. In this testing process, self-report assessments are used to screen the elderly person's driving ability in the phase of applying for testing through the local community to verify a potential need for specialized tests [37]. Based on this need, various countries use self-assessment tools to monitor the population of elderly drivers [12,13]. Likewise, the SAFE-DR was developed in South Korea as a self-report assessment for elderly drivers suited to the local culture [14], and our study has demonstrated the reliability and validity of a SAFE-DR test provided in an app. Compared to assessments that require meeting with a professional, self-assessment tools allow individuals to test their own abilities, which has the advantage of reducing stress and readily identifying potential problems in a broad population of elderly drivers [23]. In addition, mobile-based services can provide the elderly with various benefits such as access to diverse information, convenience, and reduced social isolation [16]. Therefore, this study's app-based SAFE-DR test can be used for South Korean elderly drivers both to conveniently test their own driving ability and determine the need for further testing.

Although services provided in a mobile environment have the advantage of easy, repeated access, if the results are not consistent each time the service is used, the user may become confused. Therefore, we analyzed the test-retest reliability when testing was repeated after 2 weeks; most items showed a very high reliability of ≥0.9, demonstrating the test's ease of access and repeated use.

Limitations

Our study has some limitations. We were unable to control certain factors during participant recruitment, resulting in a high ratio of males, a low proportion of persons aged ≥80 years, and a low proportion of elderly drivers who were not currently driving. Since we were developing an assessment to be used by South Korean elderly drivers, it was necessary to recruit a better-matched sample of participants and compare the test results according to participants' characteristics. In addition, because we only provided an app-based service for Android, it will be necessary to expand the test to other platforms to allow use by a larger population.

Key Findings

Despite these limitations, this study developed, for the first time, the SAFE-DR into an app-based self-report assessment that reflects the cultural characteristics of Korean elderly drivers. While many mobile-based services have recently been offered due to the advantages of easy access and fast information delivery, elderly persons who prefer conventional ways may have difficulty using such services [38]. This study confirmed that the results of both the app-based SAFE-DR evaluations and the paper-based assessments, which are familiar to the elderly, were consistent. Further, this study identified the factors for screening the driving abilities of Korean elderly drivers and tested the reliability for repeated use of self-assessment apps.

This assessment has implications on policies and traffic safety for Korean elderly drivers. Restrictions on the license of elderly drivers are beneficial for their traffic safety and that of the general population; however, these policies should include methods for education and preventive inspections of equipment to help maintain the licenses of elderly drivers as long as possible [39]. The app-based evaluation in this study can be used as means for preventive monitoring and education of the elderly driver population. Therefore, the assessment can help with the education and screening sections of the recent policy for renewing elderly drivers' licenses in South Korea [21]. Within these policies, this assessment can contribute to the safe continuation of driving by facilitating testing of driving ability and providing relevant information to elderly drivers in South Korea. Finally, future research should focus on expert driving ability assessment and licensing restriction systems for the policy on renewal of elderly Korean's driver's licenses.

Conclusions

In this study, we developed the SAFE-DR into an app-based self-report assessment for elderly drivers and tested its reliability and validity. In South Korea, where the aging population is rapidly increasing, the app can help elderly drivers to easily diagnose their driving skills and protect themselves from accidents while driving. It was designed to be easily used by elderly drivers and to provide essential information related to driving. Therefore, we anticipate that this assessment can contribute to safe continuation of driving by facilitating testing of driving ability and providing relevant information to elderly drivers in South Korea.

Acknowledgments

We would like to express our gratitude to all the elderly drivers who participated in this study. This work was supported by grant number NRF-2017R1C1B5076770 from the National Research Funds of Korea. The funding agency played no role in the study design; collection, analysis, and interpretation of data; writing of the report; or in the decision to submit the article for publication.

Authors' Contributions

All authors wrote the first draft of the manuscript and approved the final manuscript. SYC contributed to the study concept and design, drafting of the manuscript, and funding acquisition. HSH performed the analyses, acquired the data, and interpreted the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-Report Assessment for Elderly Driving Risk (SAFE-DR) questionnaire.
[DOCX File , 32 KB - mhealth_v9i6e25310_app1.docx]

References

1. Dickerson A, Molnar L, Bédard M, Eby D, Berg-Weger M, Choi M, et al. Transportation and aging: an updated research agenda to advance safe mobility among older adults transitioning from driving to non-driving. Gerontologist 2019 Mar 14;59(2):215-221. [doi: 10.1093/geront/gnx120] [Medline: 28958085]
2. Fraade-Blanar LA, Ebel BE, Larson EB, Sears JM, Thompson HJ, Chan KCG, et al. Cognitive decline and older driver crash risk. J Am Geriatr Soc 2018 Jul 17;66(6):1075-1081 [FREE Full text] [doi: 10.1111/jgs.15378] [Medline: 29667168]
3. Ahmed HI, Mohamed EE, Aly AM. Effect of mobility on the quality of life among older adults in geriatric home at Makkah Al-Mukarramah. Adv Life Sci Technol 2014;17:39-50 [FREE Full text]
4. Curl A, Stowe J, Cooney T, Proulx C. Giving up the keys: how driving cessation affects engagement in later life. Gerontologist 2014 Jul;54(3):423-433. [doi: 10.1093/geront/gnt037] [Medline: 23651920]
5. Doroudgar S, Chuang HM, Perry PJ, Thomas K, Bohnert K, Canedo J. Driving performance comparing older versus younger drivers. Traffic Inj Prev 2017 Jan 02;18(1):41-46. [doi: 10.1080/15389588.2016.1194980] [Medline: 27326512]
6. O'Neill D. Transport, driving and ageing. Rev Clin Gerontol 2015 Jun 16;25(2):147-158. [doi: 10.1017/s095925981500009x]
7. Siren A, Meng A. Cognitive screening of older drivers does not produce safety benefits. Accid Anal Prev 2012 Mar;45:634-638. [doi: 10.1016/j.aap.2011.09.032] [Medline: 22269551]
8. Ackerman ML, Crowe M, Vance DE, Wadley VG, Owsley C, Ball KK. The impact of feedback on self-rated driving ability and driving self-regulation among older adults. Gerontologist 2011 Jul 11;51(3):367-378 [FREE Full text] [doi: 10.1093/geront/gnq082] [Medline: 21071621]
9. Molnar LJ, Eby DW, Charlton JL, Langford J, Koppel S, Marshall S, et al. Driving avoidance by older adults: is it always self-regulation? Accid Anal Prev 2013 Aug;57:96-104. [doi: 10.1016/j.aap.2013.04.010] [Medline: 23659865]
10. Dickerson AE, Meuel DB, Ridenour CD, Cooper K. Assessment tools predicting fitness to drive in older adults: a systematic review. Am J Occup Ther 2014 Nov 01;68(6):670-680. [doi: 10.5014/ajot.2014.011833] [Medline: 25397762]
11. Molnar LJ, Eby DW, Kartje PS, St Louis RM. Increasing self-awareness among older drivers: the role of self-screening. J Safety Res 2010 Aug;41(4):367-373. [doi: 10.1016/j.jsr.2010.06.003] [Medline: 20846553]
12. McCarthy D, Mann W, Lanford D. Process and outcomes evaluation of older driver screening programs: the Assessment of Driving-Related Skills (ADReS) older-driver screening tool. Washington: National Highway Traffic Safety Administration; 2009. URL: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/811113.pdf [accessed 2021-05-31]
13. SAFER Driving: The Enhanced Driving Decisions Workbook. Ann Arbor: The University of Michigan Transportation Research Institute; 2006. URL: http://www.um-saferdriving.org/firstPage.php [accessed 2021-05-31]
14. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
15. Beer JM, Takayama L. Mobile remote presence systems for older adults: acceptance, benefits, and concerns. Proc 7th ACM/IEEE Int Conf Human-Robot Interact 2011:19-26. [doi: 10.1145/1957656.1957665]
16. Mostaghel R. Innovation and technology for the elderly: systematic literature review. J Bus Res 2016 Nov;69(11):4896-4900. [doi: 10.1016/j.jbusres.2016.04.049]
17. Tefft BC. Driver license renewal policies and fatal crash involvement rates of older drivers, United States, 1986-2011. Inj Epidemiol 2014 Dec 24;1(1):1-11 [FREE Full text] [doi: 10.1186/s40621-014-0025-0] [Medline: 27747659]
18. Langford J, Fitzharris M, Newstead S, Koppel S. Some consequences of different older driver licensing procedures in Australia. Accid Anal Prev 2004 Nov;36(6):993-1001. [doi: 10.1016/j.aap.2003.11.003] [Medline: 15350877]

19. Meuser T. License renewal policy & reporting of medically unfit drivers: descriptive review & policy recommendations. Washington: AAA Foundation for Traffic Safety; 2007. URL: https://aaafoundation.org/wp-content/uploads/2018/02/LPWorkshopRecommendationsReport.pdf [accessed 2021-05-31]
20. White S, O'Neill D. Health and relicensing policies for older drivers in the European union. Gerontology 2000 Apr 13;46(3):146-152. [doi: 10.1159/000022150] [Medline: 10754372]
21. Cho M. A study on the driver's license renewal and return policy through the recognition of the elderly's driving pattern. J Digi Converg 2018 Jul 25;16:213-222. [doi: 10.14400/JDC.2018.16.10.213]
22. Kulikov E. The social and policy predictors of driving mobility among older adults. J Aging Soc Policy 2011 Jan 30;23(1):1-18. [doi: 10.1080/08959420.2011.531991] [Medline: 21207303]
23. Siren A, Meng A. Older drivers' self-assessed driving skills, driving-related stress and self-regulation in traffic. Transp Res Traffic Psychol Behav 2013 Feb;17:88-97. [doi: 10.1016/j.trf.2012.10.004]
24. Choi S. Validity of the self-report assessment forecasting elderly driving risk (SAFE-DR) applicable to community health convergence. J Cases Inf Technol 2019;9(6):175-182. [doi: 10.22156/CS4SMB.2019.9.6.175]
25. Nyumba OT, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan 11;9(1):20-32. [doi: 10.1111/2041-210x.12860]
26. Drivers 65 plus: check your performance. American Automobile Association. 2010. URL: https://exchange.aaa.com/wp-content/uploads/2021/03/Driver-65-Plus.pdf [accessed 2021-06-01]
27. Han S, Lee J, Kim S, Cha T, Yoo D, Kim H. The standardization of the Korean Drivers 65 Plus to identify driving fitness of senior drivers. J Korean Soc Occup Ther 2018 Mar 31;26(1):15-29. [doi: 10.14519/jksot.2018.26.1.02]
28. Field A. Discovering statistics using IBM SPSS statistics. 5th Edition. London: Sage Publications; 2013.
29. 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] [Medline: 22489612]
30. Trizano-Hermosilla I, Alvarado JM. Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. Front Psychol 2016 May 26;7:769 [FREE Full text] [doi: 10.3389/fpsyg.2016.00769] [Medline: 27303333]
31. Kim H, Lee K, Kim H, Kim JH. Using mobile phones in healthcare management for the elderly. Maturitas 2014 Dec;79(4):381-388. [doi: 10.1016/j.maturitas.2014.08.013] [Medline: 25270725]
32. Comunello F, Mulargia S, Belotti F, Fernández-Ardèvol M. Older people's attitude towards mobile communication in everyday life: digital literacy and domestication processes. Hum Asp IT Aged Popul 2015 Aug;9193:439-450. [doi: 10.1007/978-3-319-20892-3_43]
33. Eby D, Molnar L, Shope J. Driving decisions workbook. Ann Arbor: University of Michigan Transportation Research Institute; 2000. URL: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/1321/94135.0001.001.pdf?sequence=2&isAllowed=y [accessed 2021-06-01]
34. Assessing fitness to drive for commercial and private vehicle drivers. Australia National Transport Commission. 2016. URL: https://austroads.com.au/__data/assets/pdf_file/0022/104197/AP-G56-17_Assessing_fitness_to_drive_2016_amended_Aug2017.pdf [accessed 2021-06-01]
35. Kim KS, Song HJ, Sohn UD. A study of need for medication education and efficient policies for elderly patients in an aging society. Yakhak Hoeji 2007;51:318-326 [FREE Full text]
36. Siren A, Haustein S. Driving licences and medical screening in old age: review of literature and European licensing policies. J Transp 2015 Mar;2(1):68-78. [doi: 10.1016/j.jth.2014.09.003]
37. Hawley CA, Smith R, Goodwin L. Road safety education for older drivers: evaluation of a classroom-based training initiative. Transp Res Part F Traffic Psychol Behav 2018 Nov;59:505-523. [doi: 10.1016/j.trf.2017.11.009]
38. Yang H, Lin S. The reasons why elderly mobile users adopt ubiquitous mobile social service. Comput Human Behav 2019 Apr;93:62-75. [doi: 10.1016/j.chb.2018.12.005]
39. Asbridge M, Desapriya E, Ogilvie R, Cartwright J, Mehrnoush V, Ishikawa T, et al. The impact of restricted driver's licenses on crash risk for older drivers: a systematic review. Transp Res Part A Policy Pract 2017 Mar;97:137-145. [doi: 10.1016/j.tra.2017.01.006]

Abbreviations

SAFE-DR: Self-Report Assessment for Elderly Driving Risk

Edited by L Buis; submitted 02.11.20; peer-reviewed by H Mehdizadeh, C Seebode; comments to author 02.02.21; revised version received 16.02.21; accepted 01.05.21; published 17.06.21.

Please cite as:
Hwang HS, Choi SY
Development of an Android-Based Self-Report Assessment for Elderly Driving Risk (SAFE-DR) App: Mixed Methods Study
JMIR Mhealth Uhealth 2021;9(6):e25310
URL: https://mhealth.jmir.org/2021/6/e25310
doi:10.2196/25310
PMID:33934068

Original Paper

Usability and Acceptance of the Embodied Conversational Agent Anne by People With Dementia and Their Caregivers: Exploratory Study in Home Environment Settings

Vera Stara[1*], PhD; Benjamin Vera[2*], MSc; Daniel Bolliger[2*], PhD; Lorena Rossi[1*], MSc; Elisa Felici[1*], PsyD; Mirko Di Rosa[3*], PhD; Michiel de Jong[4*], MSc; Susy Paolini[5*], PsyD

[1]Model of Care and New Technologies, IRCCS INRCA - National Institute of Health and Science on Ageing, Ancona, Italy

[2]iHomeLab - University of Applied Sciences & Arts, Lucerne, Switzerland

[3]Unit of Geriatric Pharmacoepidemiology and Biostatistics, IRCCS INRCA - National Institute of Health and Science on Ageing, Ancona, Italy

[4]Research Group IT Innovations in Healthcare, Windesheim University of Applied Sciences, Zwolle, Netherlands

[5]Unit of Neurology, IRCCS INRCA - National Institute of Health and Science on Ageing, Ancona, Italy

[*]all authors contributed equally

Corresponding Author:
Vera Stara, PhD
Model of Care and New Technologies
IRCCS INRCA - National Institute of Health and Science on Ageing
Via Santa Margherita, 5
Ancona, 60124
Italy
Phone: 39 071 800 4614
Fax: 39 071 35941
Email: v.stara@inrca.it

Abstract

Background: Information and communication technologies are tools that are able to support cognitive functions, monitor health and movements, provide reminders to maintain residual memory abilities, and promote social support, especially among patients with dementia. Among these technologies, embodied conversational agents (ECAs) are seen as screen-based entities designed to stimulate human face-to-face conversation skills, allowing for natural human-machine interaction. Unfortunately, the evidence that such agents deliver care benefits in supporting people affected by dementia and their caregivers has not yet been well studied. Therefore, research in this area is essential for the entire scientific community.

Objective: This study aims to evaluate the usability and acceptability of the virtual agent Anne by people living with dementia. The study is also designed to assess the ability of target users to use the system independently and receive valuable information from it.

Methods: We conducted a 4-week trial that involved 20 older adults living with dementia and 14 family caregivers in home environment settings in Italy. This study used a mixed methods approach, balancing quantitative and qualitative instruments to gather data from users. Telemetry data were also collected.

Results: Older users were particularly engaged in providing significant responses and participating in system improvements. Some of them clearly discussed how technical problems related to speech recognition had a negative impact on the intention to use, adaptiveness, usefulness, and trust. Moreover, the usability of the system achieved an encouraging score, and half of the sample recognized a role of the agent Anne. This study confirms that the quality of automatic speech recognition and synthesis is still a technical issue and has room for improvement, whereas the touch screen modality is almost stable and positively used by patients with dementia.

Conclusions: This study demonstrated the ability of target users to use the system independently in their home environment; overall, the involved participants shared good engagement with the system, approaching the virtual agents as a companion able to support memory and enjoyment needs. Therefore, this research provides data that sustain the use of ECAs as future eHealth systems that are able to address the basic and higher-level needs of people living with dementia. This specific field of research is novel and poorly discussed in the scientific community. This could be because of its novelty, yet there is an urgent need to

strengthen data, research, and innovation to accelerate the implementation of ECAs as a future method to offer nonpharmacological support to community-dwelling people with dementia.

(JMIR Mhealth Uhealth 2021;9(6):e25891) doi:10.2196/25891

KEYWORDS

dementia; older adults with dementia; embodied conversational agent; virtual personal assistant; virtual agent; virtual companion; design for older adults with dementia

Introduction

Background

The aging population around the world is growing rapidly, and dementia, as an age-dependent condition [1], has become a significant threat to global health. Worldwide, dementia could affect 47 to 132 million people by 2050, causing high impacts on individuals, families, communities, governments, and societies [2].

In this scenario, information and communication technologies, especially touch screens, are seen as tools that can support cognitive functions, monitor health and movements, provide reminders to maintain residual memory abilities, promote social support, improve communication with caregivers, and provide useful information concerning health conditions [3].

Among these technologies, embodied conversational agents (ECAs) or personal virtual assistants are seen as screen-based entities designed to stimulate human face-to-face conversational skills and thus allowing for natural human-machine interaction [4,5]. Unfortunately, the evidence that such agents are appropriate to deliver care benefits for supporting people living with dementia and their caregivers is yet to be studied. Therefore, research in this area is essential for the entire scientific community [6]. In fact, in the last 10 years, few studies have addressed the use of ECAs among older adults with dementia.

These include a tool for real-time streaming to a television of a realistic female avatar, previously programmed by a caregiver [7]. This avatar has a realistic voice, and the lips are in synchronization with its speech to ensure that its facial movements appear natural when reminders, notifications, and short dialogs with the user are used to support patients with dementia in their daily activities. In another study, a conversational agent system was shown on a computer screen in the form of an animated face resembling *a 5-year-old grandchild* [8]. This system can detect the end of the speech sound of a subject's response to a question and ask the next question. When the subject speaks, the agent reacts by automatically generating nods, mouth movements, and acknowledgments. In this specific study, the ECA is seen as an alternative way of conversing when no human conversation partner exists. The animation of a female cartoon-like character was used to develop LOUISE [9-11], displayed in an idle pose and moving its lips while speaking on either a computer screen or a television set. This ECA includes attention monitoring and interaction management to automatically determine whether a person wants to communicate. Lastly, a humanoid female character was used to investigate different affective identities found in older care home residents with Alzheimer disease [12]. The challenges of involving patients with dementia, as well as the possibility of engaging them in a home environment for a significant period, are reported in the studies as common limitations and problems.

All the examples mentioned above prove that the ECA research area in the eHealth sector is still immature [13], but this newness can open up opportunities for the future, especially for the challenge of enabling people with dementia and their caregivers to better manage their lives. To bridge this gap, this paper discusses the main findings emerging after 4 weeks in which 20 older adults with dementia and 14 family caregivers used the ECA *Anne* in home environment settings in Italy.

Objectives

This study aims to evaluate the usability and acceptability of the ECA Anne by older adults living with dementia. The study is also designed to assess the ability of target users to use the system independently and receive valuable information from it.

Methods

Overview

First, Anne was developed along the MyLifeMyWay project to enable seniors to live at home independently for as long as possible. The system was then adapted for seniors with forgetfulness, as is typical at the beginning of dementia, during the Living Well With Anne project.

The virtual character works on a Surface Pro tablet with the Microsoft Windows 10 operating system. The following languages are currently available: Dutch, English, German, Italian, and French. Anne can support people with dementia in all aspects of daily life (Figures 1 and 2): communication with the outside world, keeping track of items on the personal calendar, daily structure, medication, reading the news, and relaxation (games and music). All these functionalities and features were developed following a user-driven approach [14], with the engagement of a multidisciplinary team and the involvement of users in the requirements definition process [15,16].

Figure 1. Screenshots of current available modules showing an example layout. The layout can be different, or the module hidden or disabled, depending on user abilities.

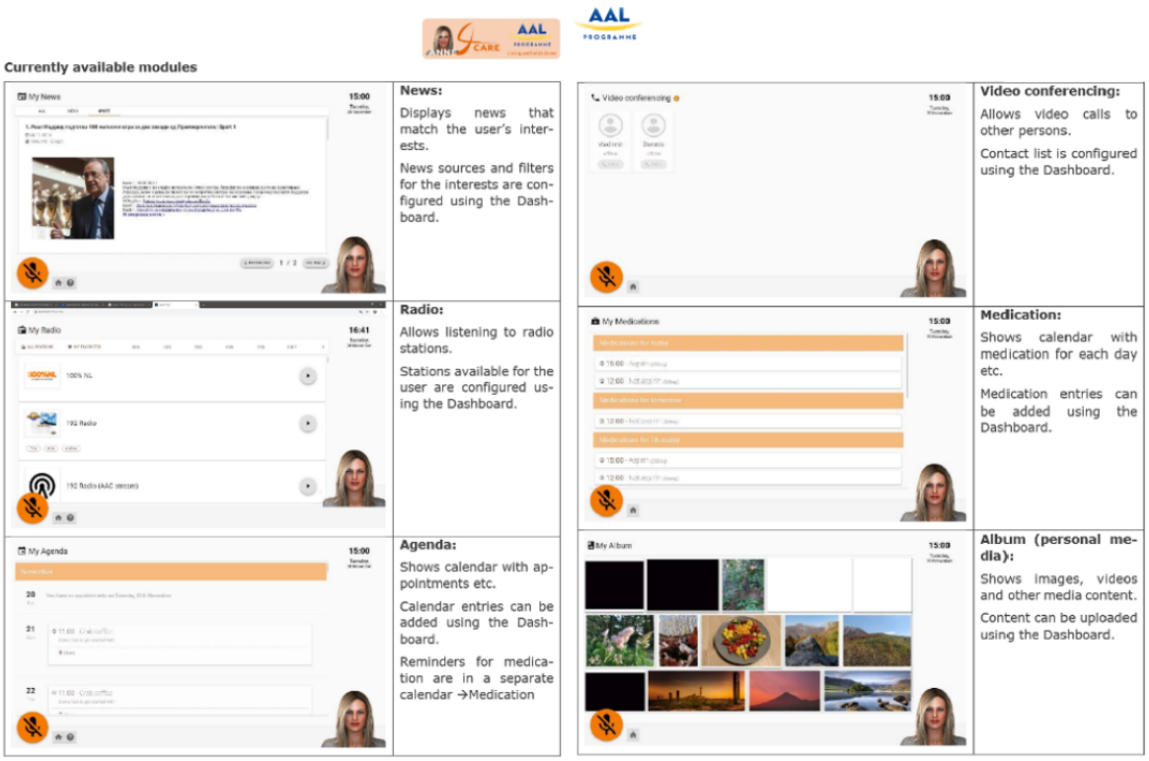
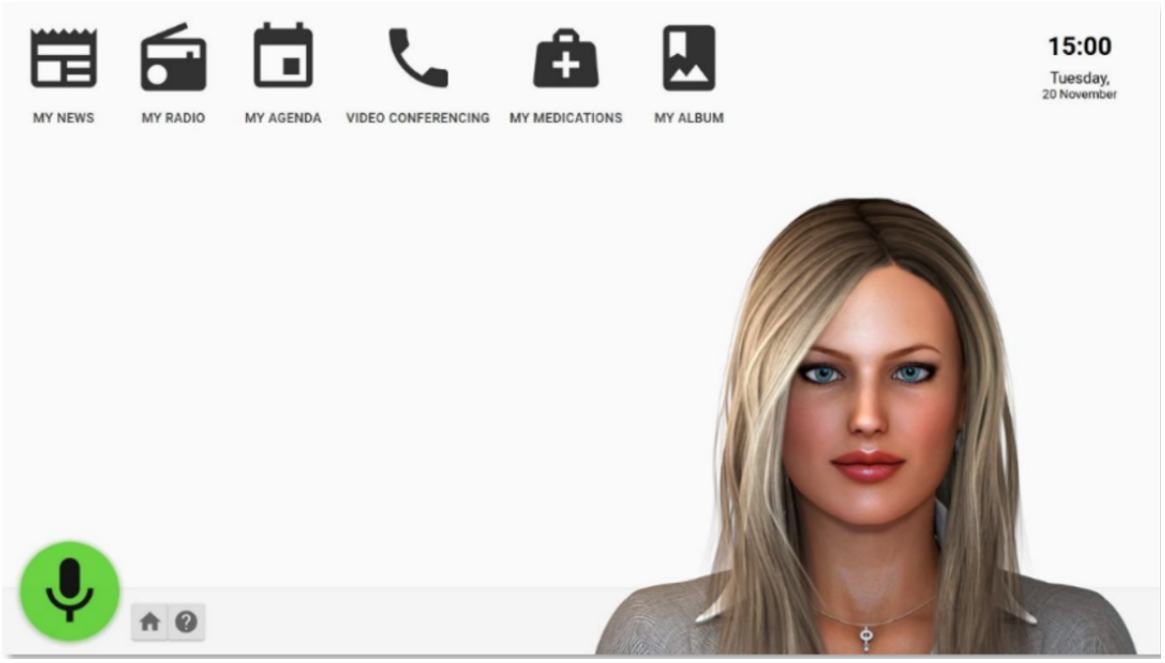


Figure 2. Home screenshots.



Users can interact with Anne through 2 channels: (1) a visual and haptic channel, via a Material–user interface graphical user interface (looking at the screen and touching the screen) and (2) an acoustical channel, via a voice user interface (listening to the avatar's voice and speaking to the avatar). Anne's voice user interface consists of automated speech recognition and text-to-speech functions. The user can always select which channel to use; that is, all commands must be accessible through touch and speech. During the requirements analysis, the clinical staff involved in the project suggested a suitable target group for the system—active and independent users able to interact with the avatar, discuss their thoughts, and express their opinions about Anne and who are competent and able to answer the protocol used for gathering data before and after the trial. Moreover, for safety reasons, individuals in the advanced stage of dementia were excluded from this study because dramatic swings in mood and behavior can be frequent, and it is not predictable how they can react in front of an unknown virtual assistant [16].

Subjects

We enrolled 20 volunteers diagnosed with dementia in the study. Inclusion and exclusion criteria for enrollment have been presented in Textboxes 1 and 2.

In total, 14 family caregivers (eg, 9 spouses and 5 sons) were involved in the study.

This study was approved by the local ethics committee, and informed written consent was obtained from all subjects.

Textbox 1. Inclusion criteria for the study.

Inclusion Criteria

- Age of 65 years or older
- Living independently
- Mini-Mental Status Examination [17] score between 24 and 27
- Ability to understand and sign the written informed consent

Textbox 2. Exclusion criteria for the study.

Exclusion Criteria

- The presence of at least one of the following criteria excluded the user from enrollment:
  - Lack of written informed consent
  - Presence of an unstable chronic condition, with a Mini-Mental Status Examination score <24
  - Presence of severe physical illness or disabilities that could be aggravated through the use of Anne

Recruitment Procedure

The enrollment and recruitment strategy was implemented in the city of Ancona at the Neurology Unit of the National Institute of Health and Science on Aging (IRCCS INRCA [Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale di Riposo e Cura per Anziani]). According to the Living Well With Anne project activities, a staff composed of 2 psychologists identified 25 eligible participants among those who had regular access to the Alzheimer daily center. Each participant was invited to test the virtual agent, Anne. The 20 users who voluntarily participated in the study and used the system were enrolled. Their caregivers were also invited to participate in the research. Five family members refused to participate in the study because of a lack of time and effort in providing care.

Study Design

The field trial ran for 4 weeks in participants' homes. The entire study was managed by skilled personnel and researchers who ensured the supervision of tests and technical assistance during the period of interaction with the system. Each enrolled subject was introduced to Anne, received general training on its correct use, and returned home with a printed user manual with step-by-step instructions and a dedicated phone number to call in case of technical problems or doubts.

This study used a mixed methods approach, balancing quantitative and qualitative instruments to gather data from users.

Users responded to the following tests at the beginning and end of the 4 weeks of use:

1. Quality of life in older adults with cognitive impairment (Quality of Life in Alzheimer Disease scale [QOL-AD]) questionnaire [18,19]. The QOL-AD questionnaire has 13 items covering physical health, energy, mood, living situations, memory, family, marriage, friends, chores, fun, money, self, and life as a whole. The assessment is scored on a 4-point Likert scale ranging from 1 (poor) to 4 (excellent), with total scores ranging from 13 to 52.
2. The Almere model [20], which is a Likert scale–based questionnaire designed primarily to measure older adults' acceptance of socially assistive robots. Almere measures the acceptance and attitudes of older adults through 12 constructs: (1) anxiety, (2) attitude toward technology (ATT), (3) facilitating conditions (FCs), (4) intention to use, (5) perceived adaptiveness (PAD), (6) perceived enjoyment (PENJ), (7) perceived ease of use (PEOU), (8) perceived sociability, (9) perceived usefulness (PU), (10) social influence, (11) social presence (SP), and (12) trust.

At the end of the period, users also responded to the questionnaires below:

1. The System Usability Scale (SUS) [21], a questionnaire that provides a quantitative measure of how usable a system is based on 10 statements rated by a 5-point Likert scale scored from 0-100, with 100 indicating perfect usability.
2. The closeness scale [22], which is a measure of self-other inclusion and relationship closeness. It was used to evaluate the closeness with the avatar at the end of the usage period.
3. Some unstructured short questions that were asked to users to record the general impression of the system (ie, role of Anne as virtual assistant, if Anne could have an impact on their well-being) and the major discomfort issues perceived during the use period.

Except for the closeness scale and unstructured short questions, family caregivers responded to the same scales. However,

because of the overburden of caregivers, time-consuming qualitative data were not gathered from them.

Each instrument was verbally administered in a face-to-face session by a trained psychologist who entered the response on a paper version of each instrument.

During the 4 weeks of use, telemetry data were collected to track every event caused by an activity of the user on the tablet. Telemetry is the process of collecting data about remote objects and sending it to a computer electronically. These activities include clicks on the touch screen or voice interaction. Moreover, the used feature types such as games or medication reminders, were also recorded. All these activities were timestamped and therefore enabled a comprehensive analysis of the user's behavior throughout time. Compared with surveys, this usage data is especially suited to detect problems and evaluate the status quo [23].

Statistical Analysis

Continuous variables were reported as mean and SD, whereas categorical variables were expressed as absolute numbers and percentages. The Almere model [20] was used as the main instrument to acquire quantitative acceptance data. Negative questions were recoded (anxiety questions 1, 2, 3, and 4; PENJ question 20; PEOU questions 21, 24, and 25; and SP question 36). The questions and constructs of the Almere model are shown in Multimedia Appendix 1.

Changing of the acceptance level from pretest to posttest for both users and caregivers were reported as mean and SD of each item of the Almere model. Two-tailed paired samples *t* tests were conducted to compare the acceptability of Anne by older adults and caregivers throughout time (pretest and posttest). Closeness and perceived relations were reported as absolute numbers and percentages. Finally, usability was measured with the SUS [21], with item values reported as mean and SD.

Results

Overview

Recruitment began in October 2019, and the trial was initiated in December 2019. Enrollment was completed in January 2020. The sample of older adults comprised 20 users (mean age 75.5 years, SD 4.2), of whom 30% (6/20) were male and 70% (14/20) were female. A large percentage of participants were married (17/20, 85%) with a medium or high level of education. Only 6 participants had previous experience using tablets for leisure activities. The caregivers (mean age 66.4 years, SD 12.6) were proportionally male and female and had a medium or high level of education. The general quality of life was in between the fair and good perception and maintained this level during the 4 weeks of the study (Table 1).

Table 1. Older adults and caregivers characteristics.

Characteristics	Older adults (n=20)	Caregivers (n=14)
Age (years), mean (SD)	75.5 (4.2)	66.4 (12.6)
Gender, n (%)		
Male	6 (30)	6 (43)
Female	14 (70)	8 (57)
Marital status, n (%)		
Married	17 (85)	11 (79)
Full-time relationship	0 (0)	1 (7)
Separated	0 (0)	0 (0)
Divorced	1 (5)	0 (0)
Single	1 (5)	2 (14)
Widowed	1 (5)	0 (0)
Education, n (%)		
No education	0 (0)	0 (0)
Primary	7 (35)	2 (14)
Secondary	4 (20)	5 (36)
Tertiary	9 (45)	7 (50)
MMSE[a], mean (SD)	25.2 (1.3)	N/A[b]
QOL-AD[c] pretest, mean (SD)	28.5 (6.6)	35.5 (5.7)
QOL-AD posttest, mean (SD)	28.9 (7.8)	34 (7.6)
Delta QOL-AD, mean (SD)	0.4 (4.6)	−1.5 (3.8)

[a]MMSE: Mini-Mental Status Examination.

[b]N/A: not applicable.

[c]QOL-AD: Quality of Life in Alzheimer Disease scale.

Acceptance

As reported in Multimedia Appendix 1, the results show that older adults became less anxious (anxiety, *P*=.007) during the 4 weeks of use. Positive changes among constructs were observed for PENJ (*P*=.04) and SP (*P*<.001). Other constructs such as ATT, PAD and FC did not change or nearly changed during the period of use. In contrast, PU (*P*=.02) and trust registered a negative change during the time of use.

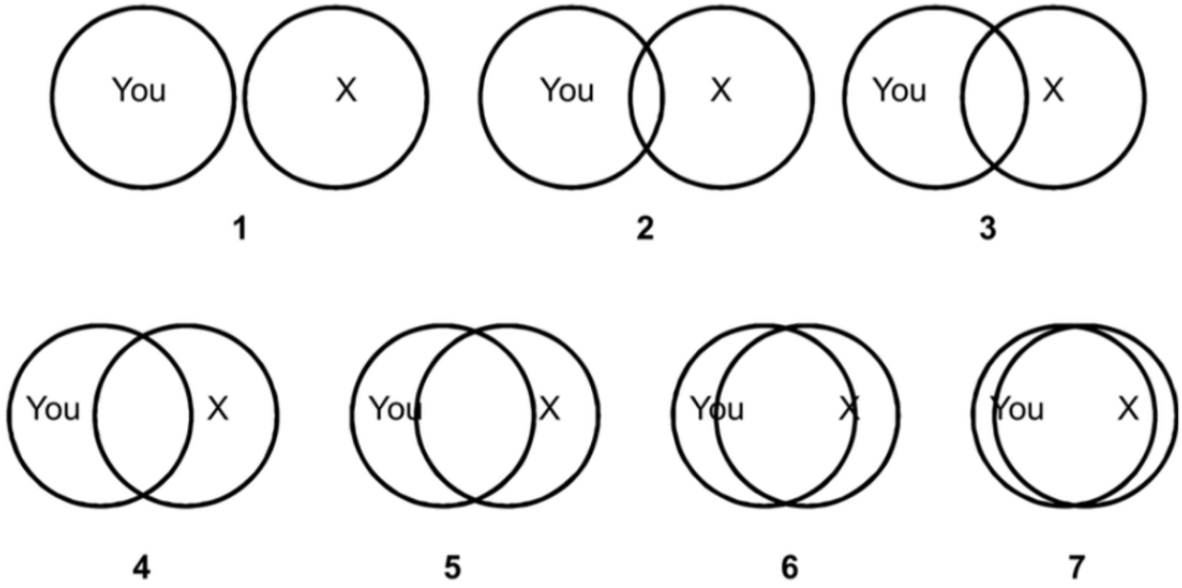
For informal caregivers, the constructs of anxiety, FC, SP, and PEOU did not change or nearly changed during interactions with Anne. The constructs that registered negative changes were ATT (*P*=.15), PAD (*P*=.02), PENJ (*P*=.005), perceived sociability (*P*=.005), PU (*P*<.001), and trust (*P*=.01).

Closeness Scale

The closeness scale aimed to assess the perceived relationship by asking respondents to evaluate their relationship with Anne. They had to select 1 of 7 pairs of increasingly overlapping circles that best described their relationship with Anne. In each pair of circles, one circle referred to the respondent and the other circle referred to Anne. A larger overlap indicated a closer relationship. For the analysis, visualization was numbered as follows: 1=no overlap, 2=little overlap, 3=some overlap, 4=equal overlap, 5=strong overlap, 6=very strong overlap, and 7=almost total overlap (Figure 3).

Older adults visualized their relationship with Anne with some overlaps (8/20, 42%) or no overlap (6/20, 26%). Minor percentages are attributed to strong, equal, or little overlap (2/20, 11%).

Figure 3. The seven pairs of increasingly overlapping circles describing the relationship between the older users and Anne.



Usability

All participants successfully completed the SUS. The SUS is scored out of 100, with a higher score indicating greater perceived usability. Anne received a mean score of 67.1 among older adults and 71.4 among caregivers. Both scores were compared and interpreted considering the acceptable average value of 68 (SD 12.5), which was determined for a variety of products and tools, including websites and technologies, provided by Sauro and Lewis [24] after the analysis of more than 5000 user scores encompassing almost 500 studies. As the virtual assistant used in this research was a system prototype and not a product ready for the market, it reached a positive result even if the score was below the acceptable average score of 68 in the case of seniors who were Anne's primary users. From the analysis of the single items reported in Table 2, participants perceived Anne as easy to use and well integrated, thus instilling confidence during use and for the idea that people could quickly learn her major functionalities.

Table 2. System Usability Scale average scores among participants.

SUS[a] items	Older adults, mean (SD)	Caregivers, mean (SD)
SUS-1. I think that I would like to use this system frequently	3.8 (1.3)	3.2 (1)
SUS-2. I found the system unnecessary complex	1.8 (1)	1.5 (0.9)
SUS-3. I thought the system was easy to use	4.1 (1.2)	4.2 (0.9)
SUS-4. I think that I would need the support of a technical person	2.9 (1.5)	2.2 (1.4)
SUS-5. I found the various functions well integrated	3.7 (0.9)	3.6 (0.9)
SUS-6. I thought there was too much inconsistency	2.4 (1.3)	2.1 (1.4)
SUS-7. I would imagine that most people would learn quickly	3.9 (0.9)	3.8 (0.7)
SUS-8. I found the system very cumbersome	2 (1.1)	1.6 (0.9)
SUS-9. I felt very confident using the system	3.3 (1.5)	3.2 (1)
SUS-10. I needed to learn a lot of things before I could get going	2.9 (1.3)	1.9 (0.9)
SUS score	67.1 (23.3)	71.4 (17.6)

[a]SUS: System Usability Scale.

General Impression of the System and Major Discomforts

Some discomfort issues related to the speech command were pointed out in the unstructured short questions asked to users:

> *The speech command does not work well. Sometimes Anne answers other questions, and this is very frustrating, making me feel insecure. Maybe this is because of me as I have no tech skills.*

About the news service, a participant mentioned the following:

> *Anne's speech is poor, does not stop for punctuation when reading, always the same monotonous tone. Moreover, Anne only reads the title and it is not possible to listen to the full article.*

Medication and gaming functions were the most successful services:

> *I often take it at the wrong time. I was very precise in my medicine intake thanks to Anna.*
>
> *I think that having an assistant who reminds me is a great help.*
>
> *The games help me keep my mind active.*
>
> *I used to play memory games and puzzles every night. I think it is perfect to stimulate my memory.*
>
> *Games helped me feel less lonely.*

In addition, caregivers also really appreciated these functions:

> *I was glad to see my mother-in-law doing something new during the day.*

The general impression was good among older users:

> *Anne kept me company when I was bored or alone.*
>
> *I enjoy talking to Anna. Every day I say 'good morning, Anne.'*
>
> *When I was lonely, I used to talk to Anna.*

Most older adults perceived Anne as a friend (6/14, 40%) or did not perceive any role to attribute (4/14, 30%). The remaining 30% (4/14) perceived the ECA as an assistant (2/14, 15%) and secretary (2/14, 15%).

When asked if Anne was seen as a way to improve their well-being, seniors responded positively (14/20, 70%), with a slight difference detected between male (4/6, 66%) and female (10/14, 71%) users. Among these users, the sense of well-being was related to the match with memory (6/17, 30%), ability to do things for fun (3/17, 15%), and mood (2/17, 10%). Female users mostly associated to the match with memory, whereas males identified a match with mood. The remaining 30% (6/17) did not find any connection between quality of life and Anne.

Telemetry Data

Two distinct classes of events were established to meaningfully analyze telemetry data. Transition events describe events for navigating through Anne, which are mostly events caused by touching on the device's touch screen, and target events that include using Anne's actual features, such as reading the news and listening to the radio. Because of the similarity with the mouse navigation, a touch on the device's touch screen is called a click in the sequel. Let us illustrate the target and transition events in a realistic example. A user would like to read a news article. Because he just started using Anne, he navigates by mistake to the game menu, realizes that and then navigates back to the main menu. Now, he navigates to the news menu and clicks on a news article that he intended to read. Each navigation caused a transition event as well as a click on the news article, resulting in 4 transition events and 1 target event. The number of transition events required to reach a target event depends on the target event itself. However, because of the user-friendliness of Anne, a target event usually requires only 1 to 2 transition events. Moreover, we see from the example how the distinction between transition and target events helps us assess the usability and learning effects of handling Anne.

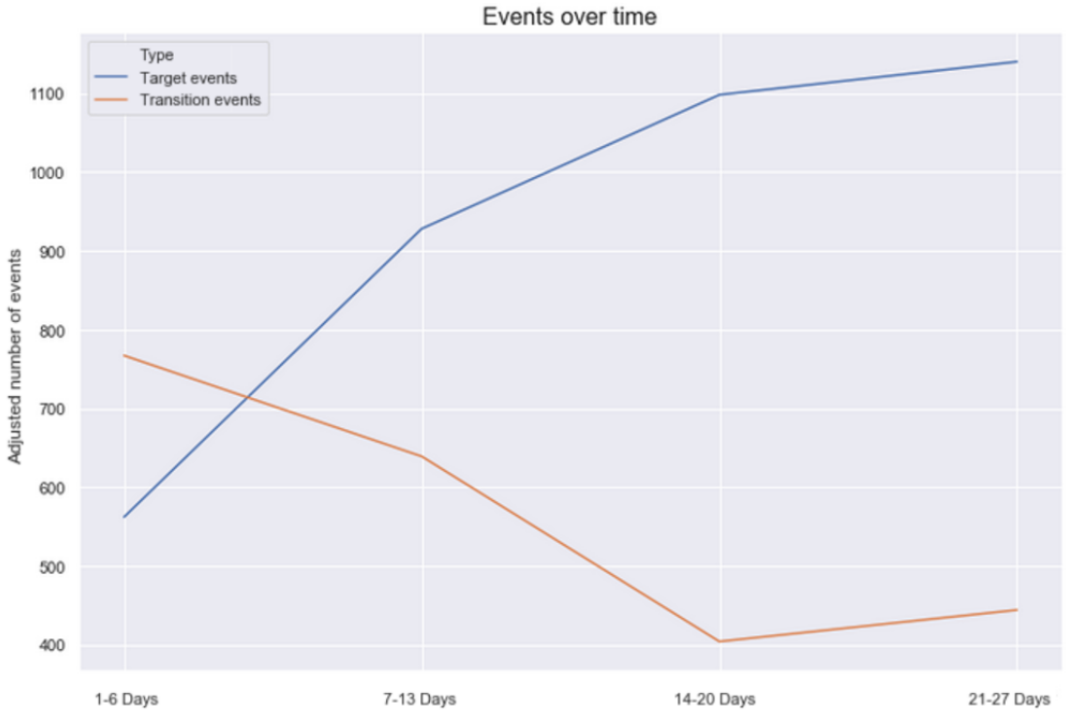
During the observed trial period, 20 users actively performed 93,299 events. The most popular feature was games (52,008 events), followed by medication (2205 events), news, and radio (with 931 and 881 events, respectively). The high number of events for games is because of Anne's design. A new event is created every time a user starts, quits, and restarts a game. These user actions are rather common in a game, and consequently, game events were very common. Nevertheless, as previously mentioned, games were a very popular feature.

From a telemetry perspective, the number of touch screen clicks before a target event can be seen as a measure of usability. Struggling users intuitively require many clicks because they try many different possibilities; hence, purposeful and quick handling is not possible. Furthermore, 80% (16/20) of the sample, that is, 16 users, only required one click on average for their target event. Thus, users seem to handle Anne very well.

Usability also has an impact on user behavior throughout time. At the beginning, a user using a new device is in an exploratory period, trying out the many features and not knowing how to handle the device very well. After some time, a user masters use and knows exactly how to handle the device. Hence, actions become very efficient and purposeful, and the efficiency in handling the device increases meaningful use. This user journey was also detected in the case of Anne.

As previously defined, transition events mainly consist of touch screen clicks and on comments provided by Anne. Thus, a large number of transition events indicate problems with handling Anne, whereas a large number of target events show actual increased purposeful activity and interest. To capture user behavior, we decomposed the active field trial time range into 4 periods. The first period lasted from day 1 to day 6, the second from day 7 to day 13, the fourth from day 14 to day 20, and the final period from day 21 to day 27. We then analyzed the number of target and transition events during these periods. As the number of users varied in these periods, the number of target and transition events was adjusted according to the respective number of users. The corresponding visualization is shown in Figure 4.

Figure 4. Target and transition events over time.



In the first and second periods, the number of transition events was 767 and 639, respectively, whereas the number of target events was 562 and 928. In these periods, the number of transition events was comparatively high, whereas the number of target events was rather low. Consequently, user exploration took place during the first 2 periods; this contrasts with the 2 subsequent periods. In the third and fourth periods, the number of target events was 1098 and 1140, respectively, whereas the number of transition events was 404 and 444, respectively. Thus, we detected a sharp increase in target events and a decrease in transition events with time. Here, the users efficiently handled Anne and had many meaningful interactions, that is, target events.

Using similar measures, weekly reports were automatically generated to summarize the usage and detect potential technical problems of users with Anne. For example, if a user stopped using Anne because he was overwhelmed. Caregivers could respond to these problems with suitable assistance and thus prevent users from becoming frustrated with Anne. This was especially helpful for the typology of users involved in this study.

Voice interactions were also assessed. Compared with 853 successful voice interactions, users still preferred the touch screen in 55,442 cases to realize their purposes. Within voice interactions, users primarily preferred the feature news in 342 cases, followed by medication in 175 cases. Touchscreen interactions were dominantly used for game features, with 52,008 cases. For example, playing a puzzle or card game really requires touch screen interactions. In this context, voice interactions are inefficient and would be very unintuitive. As games were the most popular feature by far, this had a significant impact in favor of touch screen interactions. Another possible reason for the prevalence of touch screen interaction could be that users may not be natively familiar with such a new and innovative technology; in particular, if we consider the advanced age of users.

Discussion

Principal Findings

Overall, participants involved in this study demonstrated a positive approach to Anne: after 4 weeks of use, they were less anxious about interaction with Anne and more skilled in basic functionalities, and half of them perceived a role for the ECA. None of the participants withdrew from the trial, and they all provided useful feedback to facilitate the understanding of the data gathered. On the basis of these findings, it was found that older users were particularly engaged in providing significant responses and participated in improving the system. Some of them clearly discussed how technical problems related to speech recognition had a negative impact on intention to use, adaptiveness, usefulness, and trust. This impact emerged from the qualitative data reporting discomfort issues related to speech commands. The innovative aspect of Anne, as well as of other ECAs, is that users can interact with the system through voice and by stimulating a more human-like interaction rather than just navigating with the touch screen. Unfortunately, our study confirms that the quality of automatic speech recognition and synthesis is still a technical issue and has room for improvement [7-11], whereas touch screen modality is almost stable and positively used by people living with dementia [3,25,26]. The use of voice could also be relevant for older users in general and even more so for those who, for example, have motor skill impairments such as Parkinson disease. Moreover, it could also reduce the feeling of loneliness, which is a public health concern affecting our aging society globally [27].

Despite this issue, the closeness scale [22] analysis showed that Anne was perceived as a companion able to support memory

and enjoyment needs. Anne served as a source of entertainment and as a way to handle adherence to medication plans. This is evident from telemetry data insights, which enabled the user-centric design analysis of Anne. For instance, it opens the idea for further developing actions to improve the game and medication features as well as setting new incentives for less-used features. Despite these results, the data have raised authors' awareness to the fact that ECAs could be a promising way to cope with the health and well-being of people with dementia if they are designed, developed, and assessed around and with the users. The first challenge is to target the disease process from its earliest stages and follow the person throughout the journey to foster healthy aging and improve the lives of older people, their families, and the whole community.

Our study recruited 20 users in the early stages of dementia who scored between 24 and 27 in the Mini-Mental State Examination [17], as the main objectives were to assess users' acceptance and the usability and feasibility of operating the system. In this stage, users used devices independently or, in the case of those with poor digital experience, they were still able to manage routine changes, such as introducing a new device into their daily life. In this study, users also benefited from general training on the correct use of the system, a printed user manual with step-by-step instructions, and a dedicated phone number to call for assistance in the case of technical problems or doubts. For people living with dementia, all novelties can become extremely distressing and disorientating. Therefore, becoming familiar with the technology at an early stage is fundamental, as it provides users with continuous support.

It is well known that the abilities and needs of people with dementia and those who provide care for them change along the path defined by the progression of symptoms, as does the ability to cope with them. These changing needs also mean that some technologies will be more appropriate or effective at different stages of dementia [28]. Most interventions targeting people in the early stages of dementia and their caregivers aim to support people's memory and their ability to live independently. Examples of technologies falling into this category are GPS, communication devices, and other technologies that can help mitigate memory problems (eg, medication reminders, locators, voice cues to help perform daily activities, and *dementia-friendly* versions of household gadgets) and tools promoting the self-management of health [29]. For people within moderate to severe stages of dementia (eg, with an Mini-Mental State Examination score between 20 and 9), the largest set of technologies are those that enhance safety (ie, fall detectors, motion-sensitive lights, sensors measuring room temperature and raising alarms when it gets too warm or too cold, and cooker or smoke alarms). At this stage of dementia, active use of safety technology tends to shift to the carer, whereas people with dementia often become less active users [28]. Moreover, examining feedback from the point of view of family caregivers, as in our study, or other caregivers (ie, nurses, health care professionals, or care workers who work with people with dementia), could build more awareness on how to develop effective technologies.

We cannot achieve the challenge of targeting the disease process from its earliest stages without changing the way of thinking, feeling, and acting toward people living with dementia and their complex needs. Thus, the second challenge is to support people living with dementia in doing what they need and decide, recognizing their purpose, identity, and independence. The key point is to avoid any type of stereotype about the experience of living with dementia and the opportunities that ECAs could offer to really address individuals' medical, cognitive, psychological, environmental, cultural, and social needs [30]. According to the World Health Organization global strategy on aging and health [31-33], such action could value the person's functions and needs. This means that technologies must continue to monitor health and safety as primary needs. However, it is necessary to support and maintain physical and mental capacity throughout the life course by providing opportunities for leisure and social activities to facilitate inclusion and participation, thus reducing loneliness and social isolation [34]. Supporting higher-level needs such as belonging, self-esteem, identity, and self-actualization [35] is the aim of the next generation of technologies.

Older adults are bearers of value for the design of technologies and beneficiaries of such systems. If the meaningful engagement of patients with dementia is essential for setting the future of ECAs, the two challenges depend on how much researchers, medical scientists, technology developers, and social and business innovators are ready to agree on a common vision concerning healthy aging as the way of developing and maintaining functional ability enabling well-being in old age [34-36]. In this study, the connectivity between information and communication technologies sectors, clinical staff, and research fields was ensured by adopting a user-centric design approach [14], enabling Anne's original version to be adapted to the specificities of users. Design for older adults is typically a multidimensional process involving significant time and cost in thinking, problem-solving research, iterative testing, and redesigning to meet the needs, capabilities, and limitations of users [37]. However, there are still great opportunities to be discussed and learned from the untold stories of implementing a user-centered design to create more efficient, effective, and sustainable eHealth solutions [38]. The COVID-19 pandemic highlights the significant value of digital technologies for reaching older adults, especially if frail and with multiple chronic diseases. In times of social distancing and reduced access to health services, a wide scope for innovations is covering clinical and cultural difficulties caused by the coronavirus pandemic and opening the great opportunity to increase the quality of services and access to health information.

Comparison With Previous Works and Limitations

Longer trials are needed to measure changes in user experience and familiarity with the system. This research focused on 4 weeks of interaction between 20 seniors with dementia and the agent Anne in a home setting. This framework is rare in ECA research, which mainly comprises studies on short-term interaction in a controlled environment [13] and smaller sample size enrollment [7-12]. To the best of our knowledge, this is one of the few studies on the use of ECA among people living with dementia and their caregivers. Despite these strengths, 4 weeks were not enough to evaluate a significant level of acceptability and usability, even in light of the technical

discomfort related to automatic speech recognition. Moreover, the specific Italian national context and culture could be seen as a bias and a significant limitation that does not allow for the generalization of results. Nevertheless, conducting methodologically sound scientific research in dementia care and support community is an urgent step forward [39,40]. Specifically, the possibility of running randomized control trials, enrolling a larger sample, and gathering data before and after longer interventions through robust methods remains a key challenge for the whole sector of innovation technologies [41-43]. Another key step could be to further analyze the user characteristics that match the positive acceptance of such systems and consent to better profile future customers. From a technical point of view, this study mapped how the readiness level of ECA-based interventions grew across the years, shifting from a display on standard television sets [7] and computer screens [8-12] to mobile stand-alone solutions such as in the specific case of Anne. At first, ECA functions matched the physiological, comfort, and attachment needs [7,8,12] sought, mainly to overcome memory problems, guide patients in their daily activities, and meet their need for communication and social interaction. Advancements then started to propose features such as guiding through a task, cognitive stimulation exercises, and attention management [9,10]. In comparison with these predecessors, Anne represents a step forward by offering a multipurpose tool integrating features such as reminders (personal and medication agenda), communication (video calls), information (news), and entertainment (games and music) that support users in all aspects of daily life.

Conclusions

This study aimed to evaluate the usability and acceptance of the ECA Anne by people living with dementia. It demonstrated the ability of target users to use the system independently in their home environment and receive valuable information from it. Overall, participants shared good engagement with the system, approaching the virtual agents as a companion able to support memory and enjoyment needs. Therefore, this study provides evidence for using ECAs as future eHealth systems to address the basic and higher-level needs of people living with dementia. This specific field of research is novel and poorly discussed in the scientific community. This could be because of its novelty, yet there is an urgent need to strengthen data, research, and innovation to accelerate the implementation of ECAs as a future way of offering nonpharmacological support to community-dwelling people with dementia. In our vision, this primarily means collaboration among interdisciplinary research networks, medical scientists, technology developers, and social and business innovators and the direct engagement of older adults and their formal and informal caregivers. Furthermore, sharing the strengths and weaknesses of research is fundamental for building common knowledge from previous studies. In the midst of the COVID-19 pandemic, these key points could prove significant in improving health care services. eHealth technologies have considerable challenges to overcome, but the opportunity to increase the quality of services and access to health information for users can really make a difference in these times of the pandemic.

Acknowledgments

This study is part of the Living Well With Anne project, which was cofounded by the European Union Active and Assisted Living Programme (reference #AAL-call-2016-102). This study was also partially supported by Ricerca Corrente funding from the Italian Ministry of Health, the National Funding Agency of Switzerland: State Secretariat for Education, Research and Innovation State Secretariat for Education, Research and Innovation; the Netherlands Organization for Health Research and Development (ZonMw). Virtask was the partner coordinator of the Living Well project and is the main developer of Anne under the Copyright Virtask B.V.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Almere model constructs and items.
[DOCX File , 17 KB - mhealth_v9i6e25891_app1.docx]

References

1. Irwin K, Sexton C, Daniel T, Lawlor B, Naci L. Healthy aging and dementia: two roads diverging in midlife? Front Aging Neurosci 2018;10:275 [FREE Full text] [doi: 10.3389/fnagi.2018.00275] [Medline: 30283329]
2. Alzheimer's Disease Facts and Figures. Alzheimers Dement. 2020. URL: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiF9Nbb343xAhWBT30KHeUnC1EQFnoECAYQAA&url=https%3A%2F%2Fwww.alz.org%2Falzheimers-dementia%2Ffacts-figures&usg=AOvVaw2XI3NAt3kIVz_ZFB_hLG2f [accessed 2021-06-11]
3. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, Anderberg P, Jimenez-Fernandez L, Valero-Moreno E, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253 [FREE Full text] [doi: 10.2196/17253] [Medline: 32442136]

4. Tsiourti C, Quintas J, Ben-Moussa M, Hanke S, Nijdam N, Konstantas D. The CaMeLi Framework? A Multimodal Virtual Companion for Older Adults. In: Proceedings of SAI Intelligent Systems Conference. 2016 Sep 21 Presented at: IntelliSys'16; September 21-22, 2016; London, UK p. 196-217. [doi: 10.1007/978-3-319-69266-1_10]
5. Montenegro JL, da Costa CA, da Rosa Righi R. Survey of conversational agents in health. Expert Syst Appl 2019 Sep 1;129:56-67. [doi: 10.1016/j.eswa.2019.03.054]
6. Wargnier P, Benveniste S, Jouvelot P, Rigaud A. Usability assessment of interaction management support in LOUISE, an ECA-based user interface for elders with cognitive impairment. Technol Disabil 2018 Nov 26;30(3):105-126. [doi: 10.3233/tad-180189]
7. Carrasco E, Epelde G, Moreno A, Ortiz A, Garcia I, Buiza C, et al. Natural Interaction Between Avatars and Persons With Alzheimer's Disease. In: Proceedings of the 11th international conference on Computers Helping People with Special Needs. 2008 Jul 9 Presented at: ICCHP'08; July 9-11, 2008; Linz, Austria p. 38-45. [doi: 10.1007/978-3-540-70540-6_5]
8. Yasuda K, Aoe J, Fuketa M. Development of an Agent System for Conversing with Individuals with Dementia. In: Proceedings of the 27th Annual Conference of the Japanese Society for Artificial Intelligence. 2013 Presented at: JSAI'13; June 4-7, 2013; Toyama, Japan p. 4-7 URL: https://www.ai-gakkai.or.jp/jsai2013/webprogram/2013/pdf/979.pdf
9. Wargnier P, Malaisé A, Jacquemot J. Towards Attention Monitoring of Older Adults With Cognitive Impairment During Interaction With an Embodied Conversational Agent. In: Proceedings of the 3rd IEEE International Workshop on Virtual and Augmented Assistive Technology. 2015 Mar 23 Presented at: VAAT'15; March 23, 2015; Arles, France p. 23-28 URL: https://ieeexplore.ieee.org/document/7155406 [doi: 10.1109/vaat.2015.7155406]
10. Wargnier P, Carletti G, Laurent-Corniquet Y, Benveniste S, Jouvelot P, Rigaud A. Field Evaluation With Cognitively-impaired Older Adults of Attention Management in the Embodied Conversational Agent Louise. In: Presented at 2016 IEEE International Conference on Serious Games and Applications for Health. 2016 Presented at: SeGAH'16; May 11-13, 2016; Orlando, FL, USA p. 1-8 URL: https://ieeexplore.ieee.org/document/7586282 [doi: 10.1109/segah.2016.7586282]
11. Wrobel J, Pino M, Wargnier P, Rigaud A. Robots et agents virtuels au service des personnes âgées : une revue de l'actualité en gérontechnologie. NPG Neurologie - Psychiatrie - Gériatrie 2014 Aug 22;14(82):184-193 [FREE Full text] [doi: 10.1016/j.npg.2014.02.007]
12. König A, Francis LE, Joshi J, Robillard JM, Hoey J. Qualitative study of affective identities in dementia patients for the design of cognitive assistive technologies. J Rehabil Assist Technol Eng 2017;4:1-15 [FREE Full text] [doi: 10.1177/2055668316685038] [Medline: 31186921]
13. ter Stal S, Kramer LL, Tabak M, op den Akker H, Hermens H. Design features of embodied conversational agents in eHealth: a literature review. Int J Hum-Comput Stud 2020 Jun 7;138:1-22. [doi: 10.1016/j.ijhcs.2020.102409]
14. Ergonomics of Human-system Interaction — Part 210: Human-centred Design for Interactive Systems. International Organization for Standardization. 2010 Mar. URL: https://www.iso.org/standard/52075.html [accessed 2021-06-09]
15. de Jong M, Stara V, von Döllen V, Bolliger D, Heerink M, Evers V. Users Requirements in the Design of a Virtual Agent for Patients With Dementia and Their Caregivers. In: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good. New York: ACM Press; 2018 Nov 28 Presented at: Goodtechs'18; November 28-30, 2018; Bologna, Italy p. 136-141 URL: https://dl.acm.org/doi/10.1145/3284869.3284899 [doi: 10.1145/3284869.3284899]
16. Stara V, de Jong M, Felici E, Bolliger D, Birrer E, von Döllen V. The Design Adaptation of the Virtual Assistant Anne for Moderate Dementia Patients and Their Formal Caregivers in Protected Environment Tests. In: Proceedings of the AHFE 2019 International Conference on Human Factors and Ergonomics in Healthcare and Medical Devices. 2019 Presented at: AHFE'19; July 24-28, 2019; Washington, USA p. 270-279 URL: https://link.springer.com/chapter/10.1007%2F978-3-030-20451-8_27 [doi: 10.1007/978-3-030-20451-8_27]
17. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
18. Logsdon R, Gibbons L, McCurry S, Teri L. Quality of life in Alzheimer's disease: patient and caregiver reports. J Ment Health Age 1999;5:21-32.
19. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosom Med 2002;64(3):510-519. [doi: 10.1097/00006842-200205000-00016] [Medline: 12021425]
20. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J of Soc Robotics 2010 Sep 4;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
21. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. London, England: Taylor & Francis; 1996.
22. Gächter S, Starmer C, Tufano F. Measuring the closeness of relationships: a comprehensive evaluation of the 'inclusion of the other in the self' scale. PLoS One 2015;10(6):e0129478 [FREE Full text] [doi: 10.1371/journal.pone.0129478] [Medline: 26068873]
23. Goodman E, Kuniavski M, Moed A. Observing the User Experience. A Practitioner's Guide to User Research. Burlington, Massachusetts, USA: Morgan Kaufmann; Sep 1, 2012:1-608.
24. Sauro J, Lewis J. Quantifying the User Experience: Practical Statistics for User Research. Burlington, Massachusetts, USA: Morgan Kaufmann; Jul 12, 2012.

25. Joddrell P, Astell AJ. Studies involving people with dementia and touchscreen technology: a literature review. JMIR Rehabil Assist Technol 2016 Nov 4;3(2):e10 [FREE Full text] [doi: 10.2196/rehab.5788] [Medline: 28582254]
26. Joddrell P, Astell A. Implementing accessibility settings in touchscreen apps for people living with dementia. Gerontology 2019;65(5):560-570. [doi: 10.1159/000498885] [Medline: 31018201]
27. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129 [FREE Full text] [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
28. Lorenz K, Freddolino PP, Comas-Herrera A, Knapp M, Damant J. Technology-based tools and services for people with dementia and carers: mapping technology onto the dementia care pathway. Dementia 2019 Feb;18(2):725-741. [doi: 10.1177/1471301217691617] [Medline: 28178858]
29. Evans J, Brown M, Coughlan T, Lawson G, Craven M. A Systematic Review of Dementia Focused Assistive Technology. In: International Conference on Human-Computer Interaction: Interaction Technologies. 2015 Presented at: HCI'15; August 2-7, 2015; Los Angeles, USA p. 406-417. [doi: 10.1007/978-3-319-20916-6_38]
30. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
31. World Health Organization. Global Strategy and Action Plan on Ageing and Health. Geneva, Switzeraland: World Health Organization; 2017.
32. World Health Organization. WHO clinical consortium on healthy ageing: topic focus: frailty and intrinsic capacity. In: Who Clinical Consortium on Healthy Ageing Topic Focus: Frailty and Intrinsic Capacity. Geneva, Switzerland: World Health Organization; Dec 1, 2016:1-22.
33. World Health Organization. World Report on Ageing and Health. Geneva, Switzerland: World Health Organization; 2015:-246.
34. Wang G, Marradi C, Albayrak A, van der Cammen TJ. Co-designing with people with dementia: a scoping review of involving people with dementia in design research. Maturitas 2019 Sep;127:55-63 [FREE Full text] [doi: 10.1016/j.maturitas.2019.06.003] [Medline: 31351521]
35. Koo BM, Vizer LM. Examining mobile technologies to support older adults with dementia through the lens of personhood and human needs: scoping review. JMIR Mhealth Uhealth 2019 Nov 11;7(11):e15122 [FREE Full text] [doi: 10.2196/15122] [Medline: 31710305]
36. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017 Jan 16;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
37. Czaja S, Boot W, Charness N, Rogers W. Individual differences. In: Designing for Older Adults Principles and Creative Human Factors Approaches. Boca Raton, Florida: CRC press; 2019.
38. Cornet VP, Toscos T, Bolchini D, Ghahari RR, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
39. D'Cunha NM, Nguyen D, Naumovski N, McKune A, Kellett J, Georgousopoulou E, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440 [FREE Full text] [doi: 10.1159/000500040] [Medline: 31108489]
40. Mendez MF, Joshi A, Jimenez E. Virtual reality for the assessment of frontotemporal dementia, a feasibility study. Disabil Rehabil Assist Technol 2015 Mar;10(2):160-164. [doi: 10.3109/17483107.2014.889230] [Medline: 24524440]
41. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 May 9;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]
42. Kramer LL, Stal TS, Mulder BC, de Vet E, van Velsen L. Developing embodied conversational agents for coaching people in a healthy lifestyle: scoping review. J Med Internet Res 2020 Feb 6;22(2):e14058 [FREE Full text] [doi: 10.2196/14058] [Medline: 32022693]
43. Bevilacqua R, Casaccia S, Cortellessa G, Astell A, Lattanzio F, Corsonello A, et al. Coaching through technology: a systematic review into efficacy and effectiveness for the ageing population. Int J Environ Res Public Health 2020 Aug 15;17(16) [FREE Full text] [doi: 10.3390/ijerph17165930] [Medline: 32824169]

Abbreviations

ATT: attitude toward technology
ECA: embodied conversational agent
FC: facilitating condition
IRCCS INRCA: Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale di Riposo e Cura per Anziani
PAD: perceived adaptiveness

PENJ: perceived enjoyment
PEOU: perceived ease of use
PU: perceived usefulness
QOL-AD: Quality of Life in Alzheimer Disease scale
SP: social presence
SUS: System Usability Scale

Edited by L Buis; submitted 19.11.20; peer-reviewed by K Moore, N Martin; comments to author 31.12.20; revised version received 03.02.21; accepted 10.03.21; published 19.06.21.

Please cite as:
Stara V, Vera B, Bolliger D, Rossi L, Felici E, Di Rosa M, de Jong M, Paolini S
Usability and Acceptance of the Embodied Conversational Agent Anne by People With Dementia and Their Caregivers: Exploratory Study in Home Environment Settings
JMIR Mhealth Uhealth 2021;9(6):e25891
URL: https://mhealth.jmir.org/2021/6/e25891/
doi:10.2196/25891
PMID:34170256

Original Paper

Effect of Mobile Phone Text Message Reminders on the Completion and Timely Receipt of Routine Childhood Vaccinations: Superiority Randomized Controlled Trial in Northwest Ethiopia

Zeleke Abebaw Mekonnen[1], BSc, MSc; Kassahun Alemu Gelaye[2], MPH, PhD; Martin Were[3], MD, MS; Binyam Tilahun[1], MPH, MSc, PhD, PD

[1]Department of Health Informatics, Institute of Public Health, University of Gondar, Gondar, Ethiopia

[2]Department of Epidemiology and Biostatistics, Institute of Public Health, University of Gondar, Gondar, Ethiopia

[3]Department of Biomedical Informatics, Vanderbilt Medical Center, Nashville, TN, United States

Corresponding Author:
Zeleke Abebaw Mekonnen, BSc, MSc
Department of Health Informatics, Institute of Public Health
University of Gondar
Chechela Street
Gondar, 196
Ethiopia
Phone: 251 913747092
Email: zelekeabebaw7@gmail.com

Abstract

Background: Nonattendance at vaccination appointments is a big challenge for health workers as it is difficult to track routine vaccination schedules. In Ethiopia, 3 out of 10 children have incomplete vaccination and the timely receipt of the recommended vaccines is low. Thus, innovative strategies are required to reach the last mile where mobile technology can be effectively utilized to achieve better compliance. Despite this promising technology, little is known about the role of text message–based mobile health interventions in improving the complete and timely receipt of routine childhood vaccinations in Ethiopia.

Objective: This trial aimed to determine the effect of mobile phone text message reminders on the completion and timely receipt of routine childhood vaccinations in northwest Ethiopia.

Methods: A two-arm, parallel, superiority randomized controlled trial was conducted in 9 health facilities in northwest Ethiopia. A sample size of 434 mother-infant pairs was considered in this trial. Randomization was applied in selected health facilities during enrollment with a 1:1 allocation ratio by using sealed and opaque envelopes. Participants assigned to the intervention group received mobile phone text message reminders one day before the scheduled vaccination visits. Owing to the nature of the intervention, blinding of participants was not possible. Primary outcomes of full and timely completion of vaccinations were measured objectively at 12 months. A two-sample test of proportion and log-binomial regression analyses were used to compare the outcomes between the study groups. A modified intention-to-treat analysis approach was applied and a one-tailed test was reported, considering the superiority design of the trial.

Results: A total of 426 participants were included for the analysis. We found that a higher proportion of infants in the intervention group received Penta-3 (204/213, 95.8% vs 185/213, 86.9%, respectively; *P*<.001), measles (195/213, 91.5% vs 169/213, 79.3%, respectively; *P*<.001), and full vaccination (176/213, 82.6% vs 151/213, 70.9%, respectively; *P*=.002; risk ratio 1.17, 95% lower CI 1.07) compared to infants in the usual care group. Similarly, a higher proportion of infants in the intervention group received Penta-3 (181/204, 88.7% vs 128/185, 69.2%, respectively; *P*<.001), measles (170/195, 87.1% vs 116/169, 68.6%, respectively; *P*<.001), and all scheduled vaccinations (135/213, 63.3% vs 85/213, 39.9%, respectively; *P*<.001; risk ratio 1.59, 95% lower CI 1.35) on time compared to infants in the usual care group. Of the automatically sent 852 mobile phone text messages, 764 (89.7%) were delivered successfully to the participants.

Conclusions: Mobile phone text message reminders significantly improved complete and timely receipt of all recommended vaccines. Besides, they had a significant effect in improving the timely receipt of specific vaccines. Thus, text message reminders can be used to supplement the routine immunization program in resource-limited settings. Considering different contexts, studies on the implementation challenges of mobile health interventions are recommended.

Trial Registration: Pan African Clinical Trial Registry PACTR201901533237287; https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=5839

(JMIR Mhealth Uhealth 2021;9(6):e27603) doi:10.2196/27603

KEYWORDS

mHealth; eHealth; mobile phone; text message; short message service; reminder; immunization; vaccination; Ethiopia

Introduction

Background

Vaccine-preventable diseases (VPDs) remain a common cause of childhood mortality with an estimated 3 million deaths globally each year. Vaccines have proven to be one of the most effective preventive interventions and they provide successful means for controlling VPDs [1-4]. The World Health Organization estimates that 29% of deaths in children younger than 5 years can be averted with existing vaccines [5]. Evidence also shows that being fully vaccinated is associated with 22% lower mortality in children [6]. Improving vaccination coverage is a priority for global health [7,8], and timeliness is an essential component of child vaccination that reduces susceptibility to VPDs [9]. The World Health Organization recommends that vaccines must be given within specified vaccination schedules and intervals [10] although there is no agreed upon definition of timeliness of vaccination across different countries [11]. Vaccine effectiveness also depends on the timing of its administration [12,13].

Despite the tremendous efforts to improve immunization programs, vaccination coverage and timeliness have remained suboptimal [14-16]. Globally, in 2018, coverage of the third dose of a vaccine protecting against diphtheria, tetanus, and pertussis remained at 86% [17]. Improving child vaccination coverage in low-income countries is challenging, and the vaccination coverage in Africa has been reported to be stalled at 76% [17]. Reports also show that globally, 5.9 million children were partially vaccinated in 2018 leaving many children susceptible to VPDs [18]. In low-income countries, failure to attend vaccination appointments is still a significant problem to health care providers, thereby contributing to delayed and missed vaccinations [12,19,20]. Studies in low-and-middle-income countries have also shown that despite relatively high vaccination coverage, there is a gap in the timeliness of childhood immunization [21-26].

In Ethiopia, routine childhood vaccines are administered at birth, fourth week, sixth week, tenth week, and ninth month before 1 year of age [2]. However, the completeness and timeliness of vaccination are not optimal at the national level [27,28]. The Ethiopian Demographic and Health Survey 2016 report indicates that only 39% of children received all the basic vaccinations [29]. Besides, the 2019 mini Ethiopian Demographic and Health Survey reported that the full vaccination coverage has reached 43%, with a steady rise in vaccination coverage over time [30]. In terms of vaccination completion, Ethiopia has the second largest number of incompletely vaccinated children in Africa, after Nigeria [31]. A systematic review and meta-analysis on incomplete vaccination in Ethiopia also reported that 3 out of 10 children have incomplete vaccination [32]. Similarly, studies in different parts of the country revealed that the partial vaccination coverage ranged from 17% to 45.5% [33-37]. Moreover, only a small proportion of children are vaccinated on time [28,38]. According to the national immunization coverage survey report, the prevalence of valid vaccination doses for all vaccines was only 18.6% [39]. In this study setting, the baseline assessment also indicated that full vaccination coverage was 64.3% and the timely completion for all vaccines was only 31.9% [40]. Consequently, the targets set for the different vaccines are not met [28,30].

Maintaining reductions in mortality due to VPDs relies upon continued vaccination uptake that is reliant on subsequent attendance for each vaccination appointment. However, the frequently mentioned reasons for missed and delayed vaccination in children are related to the difficulty in tracking the vaccination appointments as scheduled [13,19,28,38,41]. Forgetfulness and being unaware of the need to return for subsequent doses were among the common reasons for nonattendance to vaccination schedules, which are remediable through appropriate reminder services [28,37,38,40,42-47]. Studies have shown that achieving universal coverage with all recommended vaccines requires tailored strategies that address barriers to vaccination [15,16,28,48-50]. Thus, innovative strategies are required to reach the last mile where mobile technology can be effectively utilized to augment the immunization program. Globally, the rapid development of mobile technology has created new ways for addressing public health challenges where mobile phones are gradually becoming an integral part of health care services worldwide [51]. Mobile health (mHealth) is the use of mobile phone technology to deliver health care [52]. Among the mobile phone features, SMS remains one of the most popular forms of mobile communication with the potential to reach a large number of individuals at relatively low cost [53-55]. The World Health Organization also reported that SMS is the most common mobile phone feature used for appointment reminders to strengthen the health care system [52].

Different studies have shown that text message reminders are promising in improving attendance at health facilities and health outcomes [47,56-61]. Several studies have suggested that parental reminders for vaccination should include modern technologies such as mobile phone text message reminders [62-65]. However, the effectiveness of mobile phone–based text message reminders in improving the immunization program varies between settings. We identified studies that examined mobile phone text message reminders for vaccination, some of which found it effective in increasing vaccination uptake [44,62,66-72] while the other studies found no statistically significant effect on child vaccination [73-77]. Thus, the

scientific body of evidence on the current practice in low-income countries is still limited [14,63,72,77-79].

Objectives

In Ethiopia, mobile phone technology access is expanding rapidly [80,81]. Despite this promising and fast-growing technology, little is known about the possible role of mHealth interventions in improving the immunization program in the Ethiopian local context. Hence, this trial was conducted to assess the effect of mobile phone text message reminders in improving child vaccination service uptake in Gondar city, northwest Ethiopia.

Methods

Trial Design

This study applied a two-arm, parallel, superiority, individually randomized controlled trial design with a 1:1 allocation ratio. This trial followed a published study protocol that details the methods and approaches implemented in the trial [82]. This randomized controlled trial was conducted in the public health facilities (8 health centers and 1 comprehensive specialized hospital) of Gondar city, northwest Ethiopia between May 2019 and June 2020. This city has an estimated total population of 390,644 of which 12,149 are younger than 1 year [83]. The baseline assessment in the study area indicated that 78.9% (360/456) of the mothers had the intention to use mobile phone text message reminders for child vaccination. The assessment also pointed out that among mothers who owned a mobile phone, 91.0% (415/456) can read text messages, which makes the trial feasible in the study setting [84].

Participants

For this particular study, eligible mother-infant pairs from the University of Gondar Comprehensive Specialized Hospital and all the 8 health centers were included. The eligibility criteria for the study participant selection are shown in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the participants in this study.

Inclusion Criteria

- Mothers who have an infant who took the Bacillus Calmette-Guérin vaccination up to 4 weeks of age
- For twin infants presented for vaccination, the younger infant was included
- Mothers aged 18 years and older
- Mothers who had a working mobile phone
- Resided in the study area at least for 6 months (permanent residents)
- Mothers who were willing to provide consent for participation in the study

Exclusion Criteria

- Infants who already received vaccinations other than Bacillus Calmette-Guérin or polio zero vaccines
- Mothers who could not read mobile phone text messages in Amharic or English languages
- Mothers who had no mobile network access in their house/compound
- Mothers who planned to relocate out of the study area during the study follow-up period

During enrollment, when more than one eligible mother-infant pair from the same house/compound presented to the health facility for vaccination, only one infant (younger) was included to reduce the risk of information pollution among study participants. The recruitment of the study participants started on May 1, 2019 and ended on June 26, 2019. Study participants were followed up for 12 months.

Intervention

Participants assigned to the intervention group received the routine vaccination appointment reminder and additional mobile phone text message reminders one day before the scheduled vaccination visits on the sixth week, tenth week, fourteenth week, and ninth month after childbirth. Participants in the usual care group used vaccination cards and were informed of the due date of the next vaccination schedule verbally by health care providers working in health facilities during the facility visit. For the intervention group, a computerized text message reminder system was designed and developed for this particular study considering the local context by the eHealthLab Ethiopia research team [85]. The computerized text message reminder system has 2 components: a web-based application for client registration and automatic reminder scheduling (Multimedia Appendix 1) and a mobile SMS app for sending appointment reminders (Multimedia Appendix 2). The mobile phone text message reminders were designed and developed based on literature search, consultation with experts, and considering users' preferences. Based on the findings from baseline assessments in the study area on mothers' intentions and preferences to use text message reminders [84], mobile phone text messages were developed and delivered to users both in Amharic and English languages (Multimedia Appendix 2). The detailed development process of the automated reminder system and its function is described in a separate publication [85].

Outcomes

Full vaccination coverage and on-time full vaccination coverage were the primary outcomes for this trial. The secondary outcomes included coverage and timeliness for specific vaccine doses. In this trial, an infant was defined as having achieved

full (complete) vaccination once he/she had received all the recommended vaccines included in the national immunization schedule, namely, a dose of Bacillus Calmette-Guérin, 3 doses of oral polio vaccine, 3 doses each of pentavalent and pneumococcal conjugate vaccine, 1 dose of inactivated polio vaccine, 2 doses of rotavirus vaccine, and 1 dose of measles vaccine by the age of 12 months [2,3,29,36]. Further, the timeliness of vaccine administration (age-appropriate vaccination) was determined based on the routine schedule of Ethiopia's national immunization program–recommended age of vaccination. Accordingly, on-time vaccination for specific vaccines was defined as a specific vaccine dose administered within 4 days before the due date [86-89] and within 4 weeks (inclusive of the due date) after the recommended age specified in the national immunization schedule [9,10,41,87-91]. The denominator for the timeliness of specific vaccines was considered from those infants vaccinated for that specific vaccine [9,10,41,87-91]. In addition, on-time full vaccination was defined as all vaccine doses administered within 4 days before the due date [86-89] and within 4 weeks (inclusive of the due date) after the recommended age specified in the national immunization schedule [19,41,87,90-93]. Otherwise, it was considered as not fully vaccinated on time if at least one vaccine dose was given early, late, or missed at all at the end of the follow-up. The denominator for on-time full vaccination considered all infants who were enrolled for this trial [19,41,87,90-93].

Study End Points

The effect of mobile phone text message reminders was measured by comparing the difference in the vaccination coverage and timely receipt of vaccination between the intervention and usual care groups at 6 weeks, 10 weeks, 14 weeks, 9 months, and 12 months of age.

Sample Size Determination

Before the commencement of the randomized controlled trial, a baseline assessment was conducted in the study area to obtain recent estimates on the complete and timely vaccination status of children [40]. The sample size required for this trial was determined using the STATA statistical software (Release 14. College Station, TX: StataCorp LP). The sample size was calculated for both vaccination completeness and timeliness independently. Finally, the larger sample size was considered for this trial.

Sample Size Determination Using Full Vaccination Coverage

The sample size calculation for full vaccination assumed a power of 90%, a significance level (α) of 5% for superiority design, a difference of 15% between the intervention and usual care groups from the baseline full vaccination coverage of 64.3% [40] to 79.3%, and 20% for lost to follow-up with intervention:control ratio of 1:1. With these assumptions, the sample size required for both study groups was 366 (183 in each group).

Sample Size Determination Using On-Time Full Vaccination Coverage

A sample size of 434 mother-infant pairs (217 in each of the intervention and usual care groups) was required to detect a difference of 15% between the intervention and usual care groups from the baseline on-time full vaccination coverage of 31.9% [40] to 46.9%, 90% power, significance level (α) of 5% for superiority design, and accounting 20% for lost to follow-up. Therefore, for the 2 arms, the final sample size of the trial was the larger sample size of 434 (217 in the intervention group and 217 in the usual care group).

Randomization

The units of randomization were mother-infant pairs randomized in one of the two study arms. All mother-infant pairs who were eligible and gave informed written consent for participation were randomly assigned to either of the study groups by using a simple randomization technique. Randomization was applied in the selected health facilities during enrollment with a 1:1 allocation ratio by using sealed and opaque envelopes within each health facility separately.

Allocation Concealment

To prevent foreknowledge of intervention assignment, we used identical and small-sized sealed opaque envelopes for the random sequence generation. Random sequence generation was ensured by a research assistant ahead of the study subject's enrollment with which study arms were marked on paper and folded to fit the envelope. Finally, the sealed envelopes prepared for both the intervention and usual care groups were combined for each health facility and shuffled so that the allocation sequence remained concealed. When the mother-infant pairs were found to be eligible for the trial, they were assigned to the intervention or usual care group by randomly picking up a closed envelope.

Blinding

Owing to the nature of the intervention, blinding of the study participants was not possible. Health workers who administered the vaccines and recorded the vaccination status of the infants were blinded to study group allocations. Outcome assessors were also blinded to the intervention allocation.

Implementation

This trial was implemented as per the a priori published protocol [82]. Initially, the eligibility of the mother-infant pairs presenting for Bacillus Calmette-Guérin vaccination of an infant at the vaccination units of the included health facilities was assessed. During enrollment, informed written consent was obtained from each eligible study participant. Enrollment and assignment of the intervention to the study participants were made by trained data collectors. In addition, baseline information was collected from both intervention and usual care groups on sociodemographic characteristics, health care service use, and mobile phone use characteristics by using validated data collection tools. All mothers assigned to the intervention group were assigned a unique code and provided orientation on how to read mobile phone text messages received from the automated reminder system. Thereafter, participants assigned to the

intervention group received mobile phone text message reminders sent from the automated text message reminder system one day before the due date of the scheduled vaccinations. A research assistant was assigned to manage the automated reminder system throughout the follow-up period. Data for primary and secondary outcomes were collected from written vaccination records found in the health facility's expanded program on immunization (EPI) registers by trained outcome assessors. During the follow-up, data on the vaccination status of infants were collected regularly. Validated data collection tools were used to collect outcome data [40].

Statistical Analysis

The baseline data collected from the study participants and the outcome data collected from EPI registers were entered into the EpiData version 3.1 software (EpiData Association). Finally, the data were exported to the STATA statistical software for analysis. Analyses were done with the modified intention-to-treat analysis principle at the participant level so that all randomized participants with available outcome data were included for the analysis, regardless of the degree of exposure to study intervention [94]. Initially, descriptive statistics were computed. The household wealth index was created by principal component analysis, including variables on asset ownership, housing characteristics, ownership of animals, and farming [95]. For assessing the comparability of the data on the baseline characteristics of the study participants randomized to the intervention and usual care groups, chi-square tests were performed. Proportions were used to present the vaccination coverage and timeliness for each vaccine. To assess the changes in the vaccination status between the 2 study groups, absolute differences in the proportions with corresponding 95% CIs were computed. A two-sample test of proportion was applied to assess the statistical significance of the absolute differences. Further, risk ratios (RRs) for primary outcomes were assessed and compared between the intervention and the usual care groups by using log-binomial regression analysis [96]. As an additional analysis, post hoc subgroup analysis was also considered to assess the heterogeneity of the treatment effects by subgroups of sociodemographic characteristics. For the subgroup analysis, the log-binomial model was applied and interaction terms were tested accordingly to explore potential effect modification. Taking into account the superiority design of the trial, one-tailed test was performed and reported [97]. The effect was expressed as RR with a corresponding 95% lower confidence interval. A significance level of .05 was considered for this study. Finally, a post hoc study power analysis was conducted. The conduct, analysis, and reporting of results were done in accordance with the CONSORT updated guidelines for reporting parallel group randomized trials. In addition, the trial was reported in accordance with the CONSORT-EHEALTH checklist (Multimedia Appendix 3) [98].

Ethics and Confidentiality

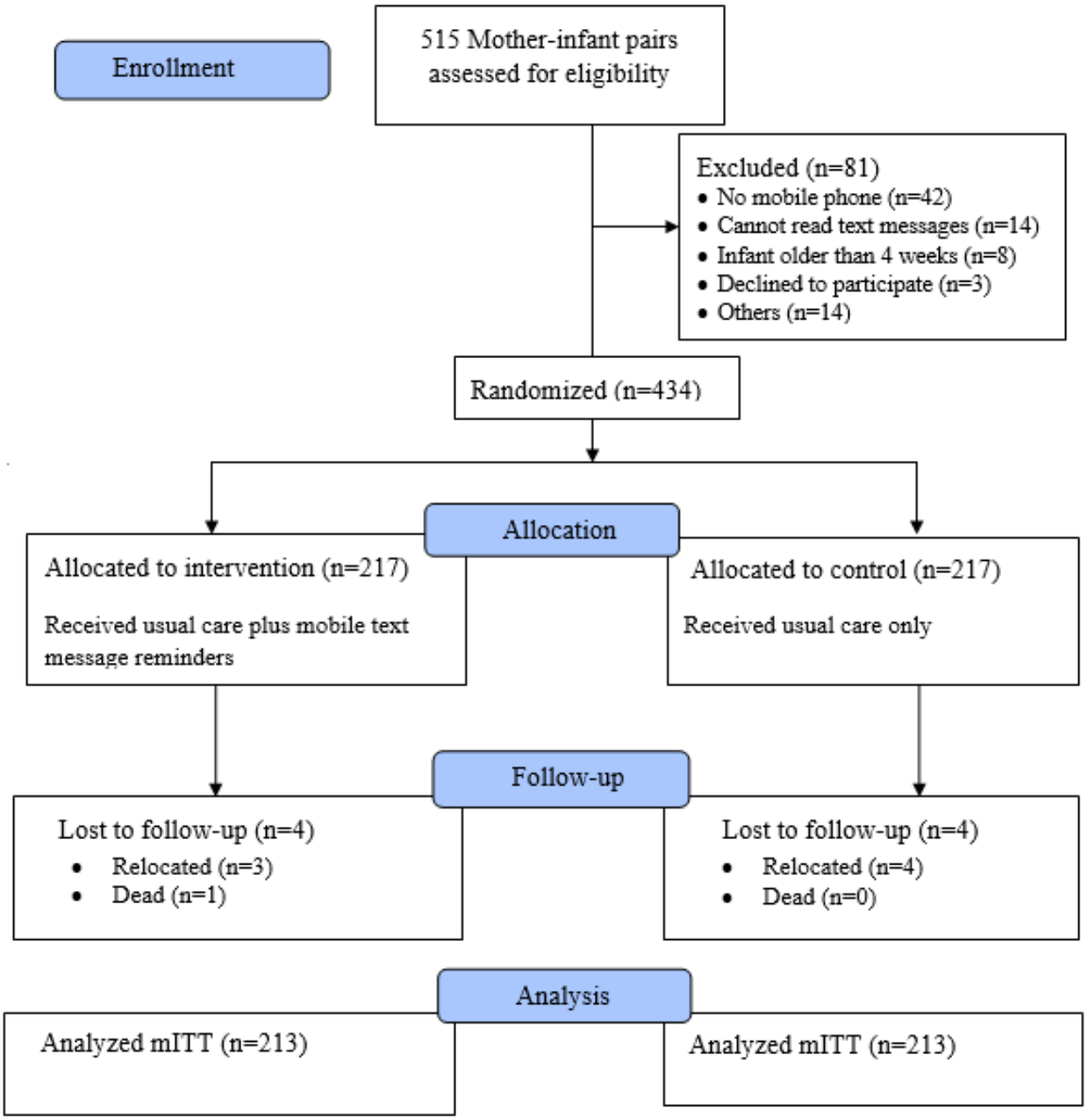
This study obtained ethical approval from the University of Gondar Institutional Ethical Review Board Reference O/V/P/RCS/05/060/2018. The purpose and procedures of the trial were explained to all the study participants. Accordingly, informed written consent was sought from all the study participants during enrollment. Confidentiality and anonymity of the information obtained were maintained at all levels of data handling. Before the commencement of the study, permission was obtained from all the required administrative levels, including the Amhara Public Health Institute and Ethio telecom.

Results

Enrollment of the Study Participants

Between May 1, 2019 and June 26, 2019, 515 mother-infant pairs were assessed for eligibility. Accordingly, 434 eligible mother-infant pairs were enrolled, with 81 excluded for not meeting the eligibility criteria (Figure 1). Of the 434 mothers enrolled, 217 were assigned to the intervention group and 217 to the usual care group. Seven participants relocated from the study area and there was 1 natural death of an infant during the study period, which actually had no relation with the outcome of interest. Thus, the analytic sample consisted of 426 mother-infant pairs who completed a 12-month follow-up in both groups with the modified intention-to-treat analysis principle (Figure 1).

Figure 1. The CONSORT flow diagram of study participant enrollment, randomization, allocation, and analysis for the trial in Gondar city, northwest Ethiopia in 2020 (N=426). mITT: modified intention-to-treat.



Characteristics of the Study Participants

The baseline characteristics of both the study groups are shown in Table 1. To validate the randomization, the characteristics of the mothers and infants of the usual care and intervention groups were examined. At baseline, the characteristics of the study participants were similar between usual care and intervention groups. The chi-square test for the baseline characteristics indicated that there was no statistically significant difference between the usual care and intervention groups at the start of the trial (Table 1).

Table 1. Baseline characteristics of the study participants enrolled in the trial in Gondar city, northwest Ethiopia in 2020 (N=426).

Characteristics	Intervention group (n=213), n (%)	Usual care group (n=213), n (%)	*P* value
Mother's age (years)			.42
≤24	64 (30.1)	71 (33.3)	
25-34	123 (57.7)	110 (51.7)	
≥35	26 (12.2)	32 (15.1)	
Marital status			.86
Currently married	195 (91.5)	194 (91.1)	
Currently not married	18 (8.5)	19 (8.9)	
Religion			.44
Orthodox	180 (84.5)	181 (84.9)	
Muslim	28 (13.1)	23 (10.8)	
Others	5 (2.4)	9 (4.3)	
Mother's education			.55
No formal education	18 (8.4)	23 (10.8)	
Primary	65 (30.5)	70 (32.9)	
Secondary and above	130 (61.1)	120 (56.3)	
Mother's occupation			.12
Housewife	105 (49.3)	100 (46.9)	
Employed	48 (22.5)	33 (15.5)	
Merchant	38 (17.9)	49 (23.0)	
Others	22 (10.3)	31 (14.6)	
Sex of infant			.77
Female	107 (50.2)	110 (51.6)	
Male	106 (49.8)	103 (48.4)	
Birth order			.16
First	81 (38.1)	95 (44.6)	
Second or later	132 (61.9)	118 (55.4)	
Residence			.53
Rural	14 (6.6)	11 (5.2)	
Urban	199 (93.4)	202 (94.8)	
Household wealth index			.30
Poor	67 (31.5)	75 (35.2)	
Middle	79 (37.0)	64 (30.1)	
Rich	67 (31.5)	74 (34.7)	
Family size			.31
<5 members	143 (67.1)	133 (62.4)	
≥5 members	70 (32.9)	80 (37.6)	
Distance to health facility			.40
<15 minutes	96 (45.1)	109 (51.2)	
15-30 minutes	82 (38.5)	70 (32.9)	
>30 minutes	35 (16.4)	34 (15.9)	
Duration of mobile phone use			.22
≤2 years	37 (17.4)	47 (22.1)	

Characteristics	Intervention group (n=213), n (%)	Usual care group (n=213), n (%)	*P* value
>2 years	176 (82.6)	166 (77.9)	
Mobile phone type			.24
Regular/standard	97 (45.5)	109 (51.2)	
Smart	116 (54.5)	104 (48.8)	

Effect of Mobile Phone Text Message Reminders on Full and Timely Completion of Vaccination

In this trial, the proportion of infants who completed the 12-month vaccination series in the intervention group was significantly higher than that of infants who completed the 12-month vaccination series in the usual care group (176/213, 82.6% vs 151/213, 70.9%, respectively; one-tailed *P*=.002). Similarly, comparisons in the timeliness of vaccination among the study groups revealed that timely completion of vaccination in the intervention group was significantly higher than that in the usual care group (135/213, 63.3% vs 85/213, 39.9%, respectively; one-tailed *P*<.001) (Table 2). Additionally, the timeliness of vaccination was compared between fully vaccinated infants in the intervention (n=176) and fully vaccinated infants in the usual care group (n=151). The findings showed that of those fully vaccinated infants in both groups, 135 (76.7%) in the intervention group and 85 (56.3%) in the usual care group were fully vaccinated on time.

Table 2. Effect of mobile phone text message reminders on full and timely completion of vaccination among infants in Gondar city, northwest Ethiopia in 2020 (N=426).

Vaccination status	Intervention group (n=213),vn (%)	Usual care group (n=213), n (%)	Absolute difference (%) (95% lower CI)[a]	*P* value[b]
Full vaccination	176 (82.6)	151 (70.9)	11.7 (5.1)	.002
On-time full vaccination	135 (63.3)	85 (39.9)	23.4 (15.7)	<.001

[a]One-tailed test reported for superiority design and there is no upper bound to the absolute difference.

[b]One-tailed test *P* value reported for the superiority design of the trial.

Effect of Mobile Phone Text Message Reminders on Specific Vaccinations

The pentavalent and measles vaccination coverages were used to measure the specific vaccination coverage for the vaccine doses administered at the age of sixth week, tenth week, fourteenth week, and ninth month vaccination schedules. At the fourteenth week, Penta-3 coverage in the intervention group was significantly higher than that in the usual care group (204/213, 95.8% vs 185/213, 86.9%, respectively; one-tailed *P*<.001). At the ninth month, the proportion of infants who were vaccinated for measles in the intervention group was significantly higher than that of infants who were vaccinated for measles in the usual care group (195/213, 91.5% vs 169/213, 79.3%, respectively; one-tailed *P*<.001) (Table 3).

Table 3. Effect of mobile phone text message reminders on receipt of specific vaccines in Gondar city, northwest Ethiopia in 2020 (N=426).

Vaccination	Intervention group (n=213), n (%)	Usual care group (n=213), n (%)	Absolute difference (95% lower CI)[a]	*P* value[b]
Penta-1	210 (98.6)	203 (95.3)	3.3 (0.6)	.02
Penta-2	209 (98.1)	193 (90.6)	7.5 (3.9)	<.001
Penta-3	204 (95.8)	185 (86.9)	8.9 (4.5)	<.001
Measles	195 (91.5)	169 (79.3)	12.2 (6.7)	<.001

[a]One-tailed test reported for superiority design and there is no upper bound to the absolute difference.

[b]One-tailed test *P* value reported for the superiority design of the trial.

Effect of Mobile Phone Text Message Reminders on Timely Receipt of Specific Vaccines

The trial showed that a significantly higher proportion of infants in the intervention group received Penta-3 vaccination on time as compared to infants in the usual care group (181/204, 88.7% vs 128/185, 69.2%, respectively; one-tailed *P*<.001). Similarly, a significantly higher proportion of infants in the intervention group also received measles vaccination on time as compared to infants in the usual care group (170/195, 87.1% vs 116/169, 68.6%, respectively; one-tailed *P*<.001) (Table 4).

Table 4. Effect of mobile phone text message reminders on timely receipt of specific vaccines in Gondar city, northwest Ethiopia in 2020.

Vaccination[a]	Intervention group, n (%)	Usual care group, n (%)	Absolute difference (95% lower CI)[b]	*P* value[c]
Penta-1	193 (91.9)	164 (80.8)	11.1 (5.6)	<.001
Penta-2	189 (90.4)	149 (77.2)	13.2 (7.3)	<.001
Penta-3	181 (88.7)	128 (69.2)	19.5 (12.9)	<.001
Measles	170 (87.1)	116 (68.6)	18.5 (11.5)	<.001

[a]Timeliness for the specific vaccines was calculated from vaccinated infants in both study groups. The total number of vaccinated infants for each vaccine is shown in Table 3.

[b]One-tailed test reported for superiority design and there is no upper bound to the absolute difference.

[c]One-tailed test *P* value reported for the superiority design of the trial.

Log-Binomial Regression Analysis of the Intervention Effect on Primary Outcomes

Log-binomial regression analysis was performed to determine the effect of mobile phone text message reminders on the primary outcomes of full and timely completion of vaccination. Mothers in the intervention group were 17% more likely to fully vaccinate their infants as compared to mothers in the usual care group (RR 1.17, 95% lower CI 1.07). In addition, mothers in the intervention group were 59% more likely to timely complete all doses of vaccines for their infants as compared to mothers in the usual care group (RR 1.59, 95% lower CI 1.35) (Table 5). The potential impact of the text message reminders was also determined using attributable risk percent. The attributable proportion among those mothers who received mobile phone text message reminders indicated that 37.1% of infants' on-time full vaccination could be attributed to the text message reminders.

Table 5. Log-binomial regression analysis of the effect of mobile phone text message reminders on the primary outcomes in Gondar city, northwest Ethiopia in 2020 (N=426).

Vaccination status	Intervention group (n=213), n (%)	Usual care group (n=213), n (%)	Risk ratio (95% lower CI)[a]
Full vaccination	176 (82.6)	151 (70.9)	1.17 (1.07)
On-time full vaccination	135 (63.3)	85 (39.9)	1.59 (1.35)

[a]One-tailed test reported for superiority design and there is no upper bound to the risk ratio.

Subgroup Analysis of the Effect of Mobile Phone Text Message Reminders on Timely Completion of Vaccination

In this trial, subgroup analysis was performed to assess the effect of the mobile phone text message reminders within categories of subgroups of the sociodemographic characteristics and to evaluate for statistically significant subgroup differences. The interaction tests showed that there were no significant effect differences across the subgroups of the included sociodemographic variables, except for the household wealth index. With the interaction test, the effect differences across the subgroups of household wealth index were statistically significant (*P*=.01). The stratum-specific findings in the subgroup analysis by household wealth index showed that the intervention effects were significantly higher for mothers who belong to the middle or rich household wealth index group (RR 1.87, 95% lower CI 1.53). However, no significant difference was found in the mothers who belonged to the poor household wealth index group (RR 1.09, 95% lower CI 0.81) (Table 6).

Table 6. Subgroup analysis of the effect of mobile phone text message reminders on timely completion of vaccination in Gondar city, northwest Ethiopia in 2020.

Characteristics	Intervention group (n=135), n (%)	Usual care group (n=85), n (%)	Stratum-specific risk ratio (95% lower CI)[a]	*P* value[b]
Mother's age				.46
≤24 years	46 (71.9)	29 (40.8)	1.76 (1.35)	
>25 years	89 (59.7)	56 (39.4)	1.51 (1.24)	
Mother's education				.13
Below secondary	44 (53.1)	25 (26.9)	1.97 (1.42)	
Secondary and above	91 (70.0)	60 (50.0)	1.40 (1.17)	
Household wealth index				.01
Poor	32 (47.8)	33 (44)	1.09 (0.81)	
Middle or rich	103 (70.6)	52 (37.7)	1.87 (1.53)	
Family size				.46
<5 members	89 (62.2)	55 (41.4)	1.51 (1.23)	
≥5 members	46 (65.7)	30 (37.5)	1.75 (1.33)	
Distance to health facility				.61
<15 minutes	63 (65.6)	47 (43.1)	1.52 (1.22)	
≥15 minutes	72 (61.5)	38 (36.5)	1.68 (1.32)	
Mobile phone type				.39
Regular/standard	55 (56.7)	43 (39.4)	1.44 (1.13)	
Smart	80 (68.9)	42 (40.4)	1.71 (1.37)	

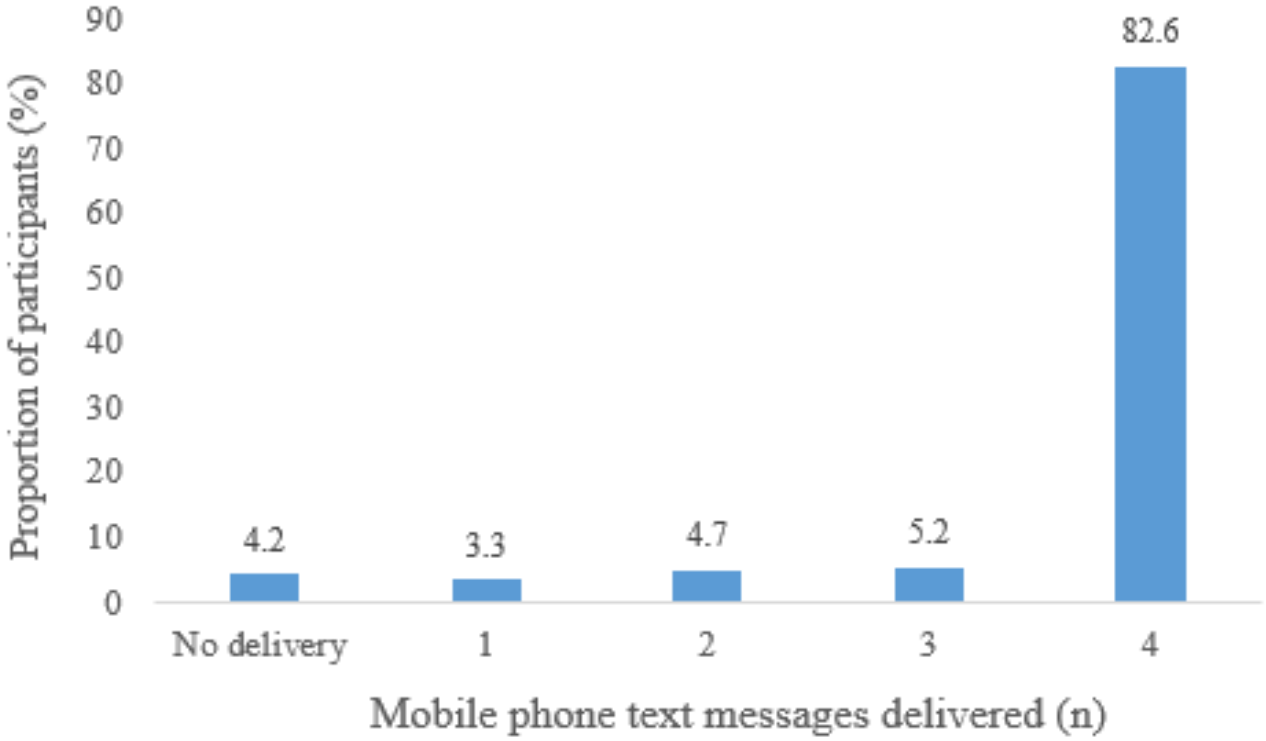
[a]One-tailed test reported for superiority design and there is no upper bound to the risk ratio.

[b]*P* value for the interaction term between intervention and baseline sociodemographic characteristics.

Mobile Phone Text Message Delivery and Reading Status

For the 213 mothers in the intervention group, 852 mobile phone text messages were dispatched successfully from the automated system for the subsequent 4 vaccination appointments. In addition, the actual delivery status of the text message to the participants was confirmed by Ethio telecom. Out of the automatically sent 852 mobile phone text messages, 764 (89.7%) were delivered successfully to the study participants. Among the 213 participants, 176 (82.6%), 11 (5.2%), 10 (4.7%), and 7 (3.3%) received 4, 3, 2, and 1 text message reminder, respectively (Figure 2). At the end of the study, we conducted a telephone survey of the participants in the intervention group to assess the reading status of the text message reminders. We attempted to contact the 204 participants in the intervention group for whom at least one text message was delivered. Among the 204 participants, 192 participants responded to our phone call. Of the total 728 text messages sent to 192 participants, 635 (87.3%) mobile phone text message reminders were seen by the study participants.

Figure 2. Mobile phone text message delivery status in the intervention group (n=213).



Study Power Analysis

For this particular trial, post hoc study power was calculated using STATA 14 software. Taking into account the superiority design of the trial, the power to detect the observed difference was 89.4% for full vaccination and 99.8% for on-time full vaccination.

Discussion

Principal Findings

This trial assessed the role of implementing a locally developed mobile phone text message reminder system to improve the immunization program in the Ethiopian context. The findings of our study indicated that text message reminders have a positive significant effect on the timely completion of routine vaccinations. This study also showed that text message reminders significantly improved timely vaccination uptake in the sixth week, tenth week, fourteenth week, and ninth month after childbirth for specific vaccines. In the intervention, of the automatically sent 852 text messages, 764 (89.7%) were delivered successfully to the study participants. This trial showed that a significantly higher proportion of infants in the intervention group were fully vaccinated as compared to infants in the usual care group (one-tailed *P*=.002). Based on our previous baseline assessments in the study area, where 34.5% of the caregivers reported nonattendance at child vaccination appointments owing to forgetfulness [40], text message reminders might help to track the schedules and improve attendance to child vaccination services. The full vaccination coverage at 12 months in both the intervention and usual care groups was lower than the national targets, indicating the importance of additional interventions to strengthen the immunization program. Our findings were consistent with those reported in studies from Bangladesh [66], Nigeria [68], and Côte d'Ivoire [71], which reported that SMS text reminders improved childhood vaccination coverage. Our findings were also in line with those of other studies that observed that mobile phone–based reminders significantly improved attendance to different health service appointments [61,75,99-101]. On the contrary, community-based studies from Ethiopia [75], Kenya [67], and South Africa [76] reported that the effect of SMS-based mobile phone interventions on full vaccination coverage was not statistically significant at 5% significance level. This difference can be explained by the differences in the study settings and differences in the study populations. It is also worth highlighting that the nature of the mHealth intervention packages might explain the observed differences. A systematic review conducted in an African setting also reported that the effect of one-way SMS reminders is uncertain and only some studies have shown the impact of SMS reminders on child vaccination attendance [57].

In this trial, infants from the intervention group had higher vaccination coverage for specific vaccines as compared to the infants from the usual care group. This finding is in line with that of other studies conducted in Zimbabwe [44], Kenya [43], and the United States of America [102,103]. However, randomized clinical trials conducted in the United States of America [73], Guatemala [74], Pakistan [70], and Nigeria [68] observed a higher vaccination coverage for specific vaccines attributable to text message reminders, which are not statistically significant at 5% significance level. The differences in the effect of the text message reminders could be related to the way text messages had been developed and delivered to the study participants.

In this study, we also observed that mobile phone text message reminders have a statistically significant effect on the timely completion of childhood vaccinations (one-tailed $P<.001$). This implies that text message reminders could serve as a valuable tool to track scheduled vaccination appointments in which the text reminders sent to clients one day before the due date of the scheduled vaccinations promptly notifies them to plan and vaccinate their child on time. Studies from Bangladesh [66], Australia [92], and Nigeria [68] corroborate this finding and reported that text reminders significantly improved the timely receipt of all scheduled vaccinations. A clustered randomized controlled trial conducted in rural Kenya also reported that SMS reminders coupled with incentives significantly improved the timeliness of vaccination [67]. However, this particular study observed that one-way mobile phone text message reminders without incentives have no statistically significant effect at 5% significance level on the timely vaccination of infants [67]. Methodological differences in measuring the timeliness of vaccination, delivery of the intervention package, and the study context may explain the effect size differences. In this study, mobile phone text message reminders increased the attendance of infants who received the recommended specific vaccines on time. Studies from the United States of America [104,105] and Nigeria [68] also reported similar findings.

In the subgroup analysis, this study showed that the intervention had positive effects on the subgroups of the sociodemographic characteristics of the study participants. This means that the intervention package appears to be beneficial for almost all mothers in the intervention group, implying that the proposed intervention has the potential to improve timely vaccination in the larger community. Besides, in the subgroup analysis, household wealth index was identified as a significant effect modifier with which mothers from middle or rich households were more benefited from the text message reminder intervention. This study also pointed out that using text message reminders is not without difficulties that need to be addressed. This was evidenced by findings of other studies that reported similar barriers [73,104]. As confirmed by phone calls at the end of the study, the majority (635/728, 87.3%) of the mobile phone text message reminders were seen by participants during the study period while the remaining text messages sent to mothers were not seen by the participants. In a real-world setting, this could be a common phenomenon, and our analysis considered the modified intention-to-treat analysis principle instead of the per-protocol analysis to declare the effectiveness of the mobile phone text message reminders in the routine EPI.

Implications for Practice and Research

Our results add unique findings to the body of literature on the impact of mobile phone text message reminders on timely vaccination, which has not been studied so far in the Ethiopian context. Our findings indicated that mobile phone–based text message reminders can be added to the arsenal of other interventions as a supplement to existing immunization programs. Moreover, the results of this study could help in guiding the future adoption and implementation of mHealth interventions in low-income countries such as Ethiopia. Program managers could consider using this system to improve child immunization service uptake as the mobile phone–based reminder system for this trial was designed in Ethiopia with the local context. This study was limited to only those mothers who have mobile phones. In implementing a text messaging reminder system, the needs of those mothers who have no mobile phones need to be addressed. Mobile phone ownership may also increase as mobile services and mobile phones become increasingly ubiquitous in low-income countries. The digital health literacy of the end users might also influence the effective implementation of mHealth interventions in different contexts, which demand prior assessment in future trials. To scale up the mobile phone–based text message interventions in the national EPI of a resource-constrained setting, further studies that guide large-scale implementation are recommended.

Strengths and Limitations

This study has the following strengths. In this trial, the study participants were allocated randomly, which resulted in balanced demographic characteristics between the study groups. This study also applied allocation concealment during the enrollment of the study participants. In addition, objective measures were used for ascertaining the vaccination status as an outcome and outcome assessors were also blind to the intervention.

Our study had the following limitations. We enrolled mothers who owned a mobile phone and presented for infant vaccination in health facilities of Gondar city, which may limit the generalizability of our findings to the general population. Owing to the nature of this study, blinding of study participants was not possible. Further, the possibility of information contamination among the participants across the study groups cannot be ruled out. The automated mobile phone text message reminder system was not set to provide information on whether mothers had received and read the text message reminders or not. Hence, the reading status of the text message reminders was confirmed at the end of the study via phone calls, which might have a possibility of recall bias.

Conclusions

In conclusion, mobile phone text message reminders significantly improved complete and timely receipt of all recommended childhood vaccinations in the study setting. Moreover, text message reminders had a significant effect in improving timely receipt of specific vaccines provided in the sixth week, tenth week, fourteenth week, and ninth month after childbirth. Thus, text message reminders can be an additional tool to usual care for improving timely completion of childhood vaccinations in resource-limited settings. Locally developed mobile phone text message reminders as a new evolution may contribute to strengthening the routine immunization program and should be considered by policy makers and program managers to improve timely completion of vaccinations. Further rigorous interventional studies in different contexts with more focus on implementation challenges of mHealth interventions are recommended.

Acknowledgments

The authors would like to thank the University of Gondar, Doris Dukes Charitable Foundation project, and the German Academic Exchange Service (DAAD) program for supporting this trial. The authors are also grateful to the Amhara Public Health Institute, Gondar City Health Department, Ethio telecom, and all the study participants. This study was financially supported by the Doris Dukes Charitable Foundation project with the award number 2017187. The funder had no role in the design, data collection, analysis, and interpretation of the study findings. The data on which these findings were developed are available upon reasonable request.

Authors' Contributions

ZAM, KAG, MW, and BT conceptualized and initiated the trial. All the authors were involved in the design and implementation of this trial. The original draft was produced by ZAM and reviewed by all authors. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Client registration, client list, vaccination appointment scheduling, and vaccination reminder analysis pages of the automated reminder system.
[PNG File , 172 KB - mhealth_v9i6e27603_app1.png]

Multimedia Appendix 2
Mobile text message reminder contents and reminder message sending process.
[PNG File , 116 KB - mhealth_v9i6e27603_app2.png]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1243 KB - mhealth_v9i6e27603_app3.pdf]

References

1. National strategy for newborn and child survival in Ethiopia:2015/16-2019/20. Ministry of Health-Ethiopia. URL: https://tinyurl.com/m5a59ksn [accessed 2020-03-05]
2. National expanded program on immunization implementation guideline. Revised Edition 2015. Ministry of Health-Ethiopia. URL: https://tinyurl.com/nxhxzmjk [accessed 2020-02-17]
3. Ethiopia national expanded program on immunization: comprehensive multi-year plan 2016-2020. Ministry of Health-Ethiopia. URL: https://tinyurl.com/ekdpan49 [accessed 2020-01-13]
4. Health sector transformation plan 2015/16-2019/20. The Federal Democratic Republic of Ethiopia Ministry of Health. URL: https://www.prb.org/wp-content/uploads/2018/05/Health-Sector-Transformation-Plan-2015_16-2019_20.Ethiopia.pdf [accessed 2019-12-16]
5. WHO, UNICEF. Global immunization data. 2014. URL: http://www.who.int/immunization/%0Amonitoring_surveillance/global_immunization_data.pdf?ua51 [accessed 2019-11-14]
6. Analysis of coverage of fully immunized child (FIC), associated factors, outcomes, and impact using routinely collected population cohort data 2001-2014. GAVI The Vaccine Alliance. URL: https://tinyurl.com/2eutpdce [accessed 2019-10-09]
7. WHO recommendations for routine immunization - summary tables. World Health Organization. URL: http://www.who.int/immunization/policy/immunization_tables/en/ [accessed 2020-07-06]
8. Nshimirimana D, Mihigo R, et al. Routine immunization services in Africa: back to basics. J Vaccines Immun 2013 Jun 01;1(1):6-12. [doi: 10.14312/2053-1273.2013-2]
9. Nalley JC, Maduka O. Completeness and timeliness of immunization among children aged 12 to 23 months in South-South Nigeria. J Community Med Prim Heal Care. 2019. URL: https://www.ajol.info/index.php/jcmphc/article/view/185038 [accessed 2021-06-07]
10. Table 2: Summary of WHO position papers - recommended routine immunizations for children. World Health Organization. 2019. URL: https://www.who.int/immunization/policy/Immunization_routine_table2.pdf [accessed 2019-10-11]
11. Masters N, Wagner A, Boulton M. Vaccination timeliness and delay in low- and middle-income countries: a systematic review of the literature, 2007-2017. Hum Vaccin Immunother 2019;15(12):2790-2805 [FREE Full text] [doi: 10.1080/21645515.2019.1616503] [Medline: 31070992]

12. Fadnes L, Nankabirwa V, Sommerfelt H, Tylleskär T, Tumwine J, Engebretsen I, PROMISE-EBF Study Group. Is vaccination coverage a good indicator of age-appropriate vaccination? A prospective study from Uganda. Vaccine 2011 Apr 27;29(19):3564-3570 [FREE Full text] [doi: 10.1016/j.vaccine.2011.02.093] [Medline: 21402043]
13. Akmatov M, Mikolajczyk R. Timeliness of childhood vaccinations in 31 low and middle-income countries. J Epidemiol Community Health 2012 Jul;66(7):e14. [doi: 10.1136/jech.2010.124651] [Medline: 21551179]
14. Oyo-Ita A, Wiysonge C, Oringanje C, Nwachukwu C, Oduwole O, Meremikwu M. Interventions for improving coverage of childhood immunisation in low- and middle-income countries. Cochrane Database Syst Rev 2016 Jul 10;7:CD008145 [FREE Full text] [doi: 10.1002/14651858.CD008145.pub3] [Medline: 27394698]
15. Chard A, Gacic-Dobo M, Diallo M, Sodha SV. Routine vaccination coverage worldwide. MMWR Morb Mortal Wkly Rep. 2019. URL: https://www.cdc.gov/mmwr/volumes/69/wr/mm6945a7.htm [accessed 2020-06-25]
16. Kazungu J, Adetifa I. Crude childhood vaccination coverage in West Africa: Trends and predictors of completeness. Wellcome Open Res 2017 Feb 15;2:12 [FREE Full text] [doi: 10.12688/wellcomeopenres.10690.1] [Medline: 28459105]
17. Peck M, Gacic-Dobo M, Diallo M, Nedelec Y, Sodha S, Wallace A. Global routine vaccination coverage. MMWR Morb Mortal Wkly Rep. 2019. URL: https://www.cdc.gov/mmwr/volumes/68/wr/mm6842a1.htm [accessed 2020-08-15]
18. WHO/UNICEF. Progress and challenges with achieving universal immunization coverage. WHO/UNICEF Estimates of National Immunization Coverage. 2019. URL: https://www.who.int/publications/m/item/progress-and-challenges-with-achievinguniversal-immunization-coverage [accessed 2020-11-18]
19. Minh An DT, Lee J, Van Minh H, Trang NTH, Thu Huong NT, Nam YS, et al. Timely immunization completion among children in Vietnam from 2000 to 2011: a multilevel analysis of individual and contextual factors. Glob Health Action 2016;9:29189 [FREE Full text] [doi: 10.3402/gha.v9.29189] [Medline: 26950555]
20. Ateudjieu J, Yakum M, Goura A, Tembei A, Ingrid D, Bita'a Landry B, et al. EPI immunization coverage, timeliness and dropout rate among children in a West Cameroon health district: a cross sectional study. BMC Public Health 2020 Feb 13;20(1):228 [FREE Full text] [doi: 10.1186/s12889-020-8340-6] [Medline: 32054484]
21. Gram L, Soremekun S, ten Asbroek A, Manu A, O'Leary M, Hill Z, et al. Socio-economic determinants and inequities in coverage and timeliness of early childhood immunisation in rural Ghana. Trop Med Int Health 2014 Jul;19(7):802-811 [FREE Full text] [doi: 10.1111/tmi.12324] [Medline: 24766425]
22. Clark A, Sanderson C. Timing of children's vaccinations in 45 low-income and middle-income countries: an analysis of survey data. Lancet 2009 May 02;373(9674):1543-1549. [doi: 10.1016/S0140-6736(09)60317-2] [Medline: 19303633]
23. Barman M, Nath K, Hazarika J. Factors Affecting Timeliness of Immunization Coverage Among Children of Assam, India. Journal of Health Management 2015 Aug 24;17(3):274-284. [doi: 10.1177/0972063415589243]
24. Mbengue M, Mboup A, Ly I, Faye A, Camara F, Thiam M, et al. Vaccination coverage and immunization timeliness among children aged 12-23 months in Senegal: a Kaplan-Meier and Cox regression analysis approach. Pan Afr Med J 2017;27(Suppl 3):8 [FREE Full text] [doi: 10.11604/pamj.supp.2017.27.3.11534] [Medline: 29296143]
25. Hu Y, Chen Y, Guo J, Tang X, Shen L. Completeness and timeliness of vaccination and determinants for low and late uptake among young children in eastern China. Hum Vaccin Immunother 2014;10(5):1408-1415 [FREE Full text] [doi: 10.4161/hv.28054] [Medline: 24584000]
26. Odutola A, Afolabi M, Ogundare E, Lowe-Jallow Y, Worwui A, Okebe J, et al. Risk factors for delay in age-appropriate vaccinations among Gambian children. BMC Health Serv Res 2015 Aug 28;15:346 [FREE Full text] [doi: 10.1186/s12913-015-1015-9] [Medline: 26315547]
27. Health sector transformation plan II 2020/21-2024/25. Ministry of Health-Ethiopia. 2020. URL: https://www.moh.gov.et/ejcc/index.php/am/node/152 [accessed 2021-06-07]
28. Tilahun B, Mekonnen Z, Sharkey A, Shahabuddin A, Feletto M, Zelalem M, et al. What we know and don't know about the immunization program of Ethiopia: a scoping review of the literature. BMC Public Health 2020 Sep 07;20(1):1365 [FREE Full text] [doi: 10.1186/s12889-020-09304-1] [Medline: 32894099]
29. Ethiopia demographic and health survey. ORC Macro and Central Statistical Agency. URL: https://dhsprogram.com/pubs/pdf/FR328/FR328.pdf [accessed 2020-05-18]
30. Mini Ethiopian demographic and health survey. Ethiopian Public Health Institute and Ministry of Health-Ethiopia. URL: https://www.ephi.gov.et/images/files/Excutive-summary---mini-DHS-2019_Final-1.pdf [accessed 2020-08-20]
31. Global routine vaccination coverage, 2015. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/mmwr/volumes/65/wr/mm6545a5.htm [accessed 2020-06-05]
32. Desalew A, Semahegn A, Birhanu S, Tesfaye G. Incomplete Vaccination and Its Predictors among Children in Ethiopia: A Systematic Review and Meta-Analysis. Glob Pediatr Health 2020;7:2333794X20968681 [FREE Full text] [doi: 10.1177/2333794X20968681] [Medline: 33241080]
33. Mohammed H, Atomsa A. Assessment of Child Immunization Coverage and Associated Factors in Oromia Regional State, Eastern Ethiopia. Sci. Technol. Arts Res. J 2013 Dec 17;2(1):36-400. [doi: 10.4314/star.v2i1.98842]
34. Tadesse H, Deribew A, Woldie M. Predictors of defaulting from completion of child immunization in south Ethiopia, May 2008: a case control study. BMC Public Health 2009 May 22;9:150 [FREE Full text] [doi: 10.1186/1471-2458-9-150] [Medline: 19463164]

35. Debie A. Assessment of Fully Vaccination Coverage and Associated Factors among Children Aged 12-23 Months in Mecha District, North West Ethiopia: A Cross-Sectional Study. SJPH 2014 May 22;2(4):342-348 [FREE Full text] [doi: 10.11648/j.sjph.20140204.26]
36. Atnafu A, Andargie G, Yitayal M, Ayele T, Alemu K, Demissie G, et al. Prevalence and determinants of incomplete or not at all vaccination among children aged 12-36 months in Dabat and Gondar districts, northwest of Ethiopia: findings from the primary health care project. BMJ Open 2020 Dec 08;10(12):e041163 [FREE Full text] [doi: 10.1136/bmjopen-2020-041163] [Medline: 33293394]
37. Workina A, Seid S, Moga T. Reason for incomplete vaccination-associated factors among children aged 12-23 months in Serbo town. Adv Res J Multi-Disciplinary Discov. 2018. URL: https://www.researchgate.net/publication/332298808_Reason_For_Incomplete_Vaccination_and_Associated_Factors_Among_Children_Aged_12-23_Months_in_Serbo_Town [accessed 2021-06-07]
38. USAID. Extended program on immunization (EPI) coverage in selected Ethiopian zones: A baseline survey for L10K's routine immunization improvement initiative. The Last Ten Kilo Meters Project. 2015. URL: https://l10k.jsi.com/Resources/Docs/EPI-baseline-report.pdf [accessed 2021-06-07]
39. Habtamu T, Abebe B, Mekonnen T, Theodros GJ, Gavin G, Amha K, et al. Ye Science Admas. 2012. URL: https://www.ephi.gov.et/images/downloads/final-scientific-newsletter-2.pdf [accessed 2021-06-07]
40. Mekonnen ZA, Gelaye KA, Were MC, Tilahun B. Timely completion of vaccination and its determinants among children in northwest Ethiopia: a multilevel analysis. BMC Public Health 2020 Jun 11;20(1):908 [FREE Full text] [doi: 10.1186/s12889-020-08935-8] [Medline: 32527248]
41. Schoeps A, Ouédraogo N, Kagoné M, Sié A, Müller O, Becher H. Socio-demographic determinants of timely adherence to BCG, Penta3, measles, and complete vaccination schedule in Burkina Faso. Vaccine 2013 Dec 17;32(1):96-102. [doi: 10.1016/j.vaccine.2013.10.063] [Medline: 24183978]
42. Patel TN, Pandit NB. Why infants miss vaccination during routine immunization sessions? Study in a rural area of Anand District, Gujarat. Indian J Public Health 2011;55(4):321-323 [FREE Full text] [doi: 10.4103/0019-557X.92417] [Medline: 22298144]
43. Haji A, Lowther S, Ngan'ga Z, Gura Z, Tabu C, Sandhu H, et al. Reducing routine vaccination dropout rates: evaluating two interventions in three Kenyan districts, 2014. BMC Public Health 2016 Feb 16;16:152 [FREE Full text] [doi: 10.1186/s12889-016-2823-5] [Medline: 26880141]
44. Bangure D, Chirundu D, Gombe N, Marufu T, Mandozana G, Tshimanga M, et al. Effectiveness of short message services reminder on childhood immunization programme in Kadoma, Zimbabwe - a randomized controlled trial, 2013. BMC Public Health 2015 Feb 12;15:137 [FREE Full text] [doi: 10.1186/s12889-015-1470-6] [Medline: 25885862]
45. Mutua M, Kimani-Murage E, Ngomi N, Ravn H, Mwaniki P, Echoka E. Fully immunized child: coverage, timing and sequencing of routine immunization in an urban poor settlement in Nairobi, Kenya. Trop Med Health 2016;44:13 [FREE Full text] [doi: 10.1186/s41182-016-0013-x] [Medline: 27433132]
46. Favin M, Steinglass R, Fields R, Banerjee K, Sawhney M. Why children are not vaccinated: a review of the grey literature. Int Health 2012 Dec;4(4):229-238. [doi: 10.1016/j.inhe.2012.07.004] [Medline: 24029668]
47. Lin C, Mistry N, Boneh J, Li H, Lazebnik R. Text Message Reminders Increase Appointment Adherence in a Pediatric Clinic: A Randomized Controlled Trial. Int J Pediatr 2016;2016:8487378 [FREE Full text] [doi: 10.1155/2016/8487378] [Medline: 28127311]
48. Ketema D, Assemie M, Alamneh A, Alene M, Chane K, Alamneh Y, et al. Full vaccination coverage among children aged 12-23 months in Ethiopia: a systematic review and meta-analysis. BMC Public Health 2020 May 24;20(1):777 [FREE Full text] [doi: 10.1186/s12889-020-08940-x] [Medline: 32448220]
49. Hargreaves A, Nowak G, Frew P, Hinman A, Orenstein W, Mendel J. Adherence to timely vaccinations in the United States. Pediatrics 2020;145(3):1-13.
50. Efendi F, Pradiptasiwi D, Krisnana I, Kusumaningrum T, Kurniati A, Sampurna M, et al. Factors associated with complete immunizations coverage among Indonesian children aged 12–23 months. Children and Youth Services Review 2020 Jan;108:104651-104655. [doi: 10.1016/j.childyouth.2019.104651]
51. Ethiopian national eHealth strategy. Ministry of Health-Ethiopia. URL: https://www.who.int/goe/policies/ethiopian_ehealth_strategy2014.pdf?ua=1 [accessed 2020-09-16]
52. World Health Organization. mHealth: New horizons for health through mobile technologies. Based on the findings of the second global survey on eHealth. Global Observatory for eHealth Series. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2020-07-14]
53. Brown VB, Oluwatosin A, Ogundeji MO. Experiences, perceptions and preferences of mothers towards childhood immunization reminder/recall in Ibadan, Nigeria: a cross-sectional study. Pan Afr Med J 2015;20:243 [FREE Full text] [doi: 10.11604/pamj.2015.20.243.6019] [Medline: 27386039]
54. Higgs E, Goldberg A, Labrique A, Cook S, Schmid C, Cole C, et al. Understanding the role of mHealth and other media interventions for behavior change to enhance child survival and development in low- and middle-income countries: an evidence review. J Health Commun 2014;19 Suppl 1:164-189 [FREE Full text] [doi: 10.1080/10810730.2014.929763] [Medline: 25207452]

55. Kalan R, Wiysonge C, Ramafuthole T, Allie K, Ebrahim F. Mobile phone text messaging for improving the uptake of vaccinations: A systematic review protocol. BMJ Open. 2014. URL: https://bmjopen.bmj.com/content/4/8/e005130.info [accessed 2021-06-07]
56. Palacios-Barahona U, Hernández-Rendón M, Gaviria-Lopez A, Guerra-Herrera A, Valencia-Díaz E, Ramos-Castañeda J, et al. Effectiveness of text messages, calls and e-mail on adherence to medical appointments. Review of systematic reviews and meta-analysis. CES Med 2018;32(1):14-22. [doi: 10.21615/cesmedicina.32.1.2]
57. Linde D, Korsholm M, Katanga J, Rasch V, Lundh A, Andersen M. One-way SMS and healthcare outcomes in Africa: Systematic review of randomised trials with meta-analysis. PLoS One 2019;14(6):e0217485 [FREE Full text] [doi: 10.1371/journal.pone.0217485] [Medline: 31170176]
58. Youssef A, Alharthi H, Khaldi OA, Alnaimi F, Alsubaie N, Alfariss N. Effectiveness of text message reminders on nonattendance of outpatient clinic appointments in three different specialties: A randomized controlled trial in a Saudi Hospital. Journal of Taibah University Medical Sciences 2014 Mar;9(1):23-29. [doi: 10.1016/j.jtumed.2013.10.001]
59. Leong KC, Chen WS, Leong KW, Mastura I, Mimi O, Sheikh MA, et al. The use of text messaging to improve attendance in primary care: a randomized controlled trial. Fam Pract 2006 Dec;23(6):699-705. [doi: 10.1093/fampra/cml044] [Medline: 16916871]
60. Junod Perron N, Dao MD, Righini NC, Humair J, Broers B, Narring F, et al. Text-messaging versus telephone reminders to reduce missed appointments in an academic primary care clinic: a randomized controlled trial. BMC Health Serv Res 2013 Apr 04;13:125 [FREE Full text] [doi: 10.1186/1472-6963-13-125] [Medline: 23557331]
61. Shiferaw S, Spigt M, Tekie M, Abdullah M, Fantahun M, Dinant G. The Effects of a Locally Developed mHealth Intervention on Delivery and Postnatal Care Utilization; A Prospective Controlled Evaluation among Health Centres in Ethiopia. PLoS One 2016;11(7):e0158600 [FREE Full text] [doi: 10.1371/journal.pone.0158600] [Medline: 27383186]
62. Wakadha H, Chandir S, Were EV, Rubin A, Obor D, Levine OS, et al. The feasibility of using mobile-phone based SMS reminders and conditional cash transfers to improve timely immunization in rural Kenya. Vaccine 2013 Jan 30;31(6):987-993 [FREE Full text] [doi: 10.1016/j.vaccine.2012.11.093] [Medline: 23246258]
63. 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;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
64. Harvey H, Reissland N, Mason J. Parental reminder, recall and educational interventions to improve early childhood immunisation uptake: A systematic review and meta-analysis. Vaccine 2015 Jun 09;33(25):2862-2880 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.085] [Medline: 25944299]
65. Bangura JB, Xiao S, Qiu D, Ouyang F, Chen L. Barriers to childhood immunization in sub-Saharan Africa: A systematic review. BMC Public Health 2020 Jul 14;20(1):1108 [FREE Full text] [doi: 10.1186/s12889-020-09169-4] [Medline: 32664849]
66. Uddin MJ, Shamsuzzaman M, Horng L, Labrique A, Vasudevan L, Zeller K, et al. Use of mobile phones for improving vaccination coverage among children living in rural hard-to-reach areas and urban streets of Bangladesh. Vaccine 2016 Jan 04;34(2):276-283 [FREE Full text] [doi: 10.1016/j.vaccine.2015.11.024] [Medline: 26647290]
67. Gibson DG, Ochieng B, Kagucia EW, Were J, Hayford K, Moulton LH, et al. Mobile phone-delivered reminders and incentives to improve childhood immunisation coverage and timeliness in Kenya (M-SIMU): a cluster randomised controlled trial. Lancet Glob Health 2017 Apr;5(4):e428-e438 [FREE Full text] [doi: 10.1016/S2214-109X(17)30072-4] [Medline: 28288747]
68. Ekhaguere OA, Oluwafemi RO, Badejoko B, Oyeneyin LO, Butali A, Lowenthal ED, et al. Automated phone call and text reminders for childhood immunisations (PRIMM): a randomised controlled trial in Nigeria. BMJ Glob Health. 2019 Apr 03. URL: https://gh.bmj.com/content/4/2/e001232.info [accessed 2021-06-07]
69. Ghadieh AS, Hamadeh GN, Mahmassani DM, Lakkis NA. The effect of various types of patients' reminders on the uptake of pneumococcal vaccine in adults: A randomized controlled trial. Vaccine 2015 Oct;33(43):5868-5872. [Medline: 26232345]
70. Kazi AM, Ali M, Zubair K, Kalimuddin H, Kazi AN, Iqbal SP, et al. Effect of Mobile Phone Text Message Reminders on Routine Immunization Uptake in Pakistan: Randomized Controlled Trial. JMIR Public Health Surveill 2018 Mar 07;4(1):e20 [FREE Full text] [doi: 10.2196/publichealth.7026] [Medline: 29514773]
71. Dissieka R, Soohoo M, Janmohamed A, Doledec D. Providing mothers with mobile phone message reminders increases childhood immunization and vitamin A supplementation coverage in Côte d'Ivoire: A randomized controlled trial. J Public Health Afr 2019 May 03;10(1):1032 [FREE Full text] [doi: 10.4081/jphia.2019.1032] [Medline: 31285815]
72. Mekonnen ZA, Gelaye KA, Were MC, Gashu KD, Tilahun BC. Effect of mobile text message reminders on routine childhood vaccination: a systematic review and meta-analysis. Syst Rev 2019 Jun 28;8(1):154 [FREE Full text] [doi: 10.1186/s13643-019-1054-0] [Medline: 31253186]
73. Ahlers-Schmidt C, Chesser A, Nguyen T, Brannon J, Hart T, Williams K, et al. Feasibility of a randomized controlled trial to evaluate Text Reminders for Immunization Compliance in Kids (TRICKs). Vaccine 2012 Aug 03;30(36):5305-5309. [doi: 10.1016/j.vaccine.2012.06.058] [Medline: 22750044]

74. Domek GJ, Contreras-Roldan IL, O'Leary ST, Bull S, Furniss A, Kempe A, et al. SMS text message reminders to improve infant vaccination coverage in Guatemala: A pilot randomized controlled trial. Vaccine 2016 May 05;34(21):2437-2443 [FREE Full text] [doi: 10.1016/j.vaccine.2016.03.065] [Medline: 27026145]
75. Atnafu A, Otto K, Herbst CH. The role of mHealth intervention on maternal and child health service delivery: findings from a randomized controlled field trial in rural Ethiopia. Mhealth 2017;3:39 [FREE Full text] [doi: 10.21037/mhealth.2017.08.04] [Medline: 29184891]
76. Coleman J, Black V, Thorson AE, Eriksen J. Evaluating the effect of maternal mHealth text messages on uptake of maternal and child health care services in South Africa: a multicentre cohort intervention study. Reprod Health 2020 Oct 20;17(1):160 [FREE Full text] [doi: 10.1186/s12978-020-01017-3] [Medline: 33081795]
77. Yunusa U, Garba SN, Umar AB, Idris SH, Bello UL, Abdulrashid I, et al. Mobile phone reminders for enhancing uptake, completeness and timeliness of routine childhood immunization in low and middle income countries: A systematic review and meta-analysis. Vaccine 2021 Jan 08;39(2):209-221. [doi: 10.1016/j.vaccine.2020.11.043] [Medline: 33277058]
78. Watterson JL, Walsh J, Madeka I. Using mHealth to Improve Usage of Antenatal Care, Postnatal Care, and Immunization: A Systematic Review of the Literature. Biomed Res Int 2015;2015:153402 [FREE Full text] [doi: 10.1155/2015/153402] [Medline: 26380263]
79. WHO. Global Diffusion of eHealth: Making Universal Health Coverage Achievable. Geneva: WHO; 2016.
80. mHealth in Ethiopia. Strategies for a new framework. Vital Wave Consulting. URL: http://vitalwave.com/wp-content/uploads/2015/09/mHealth-Framework-for-Ethiopia-2011.pdf [accessed 2020-07-13]
81. Deribe K, Roda A. The potential of mobile phone technology for public health practice in Ethiopia. Ethiop J Heal Dev. 2011. URL: https://www.researchgate.net/publication/289209352_The_potential_of_mobile_phone_technology_for_public_health_practice_in_Ethiopia [accessed 2021-06-07]
82. Mekonnen ZA, Tilahun B, Alemu K, Were M. Effect of mobile phone text message reminders on improving completeness and timeliness of routine childhood vaccinations in North-West, Ethiopia: a study protocol for randomised controlled trial. BMJ Open 2019 Nov 05;9(11):e031254 [FREE Full text] [doi: 10.1136/bmjopen-2019-031254] [Medline: 31694849]
83. Gondar city administration health office report. Gondar City Health Department. 2018. URL: https://www.facebook.com/Gondar-city-adminstration-health-department-386644484864380/about [accessed 2021-06-07]
84. Mekonnen ZA, Gelaye KA, Were MC, Tilahun B. Mothers intention and preference to use mobile phone text message reminders for child vaccination in Northwest Ethiopia. BMJ Health Care Inform 2021;8(2):1-9.
85. Mekonnen ZA, Hussien FN, Tilahun BC, Gelaye KA, Mamuye AL. Development of automated text-message reminder system to improve uptake of child vaccination in Ethiopia. Online J Public Health Inform 2019 Sep 20;11(2):e15-e17 [FREE Full text] [doi: 10.5210/ojphi.v11i2.10244] [Medline: 31632609]
86. CDC. General recommendations on immunization. Recommendations of the Advisory Committee on immunization practices. MMWR Recomm Rep. 2011. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6002a1.htm [accessed 2020-08-09]
87. Adetifa IM, Karia B, Mutuku A, Bwanaali T, Makumi A, Wafula J, et al. Coverage and timeliness of vaccination and the validity of routine estimates: Insights from a vaccine registry in Kenya. Vaccine 2018 Dec 18;36(52):7965-7974 [FREE Full text] [doi: 10.1016/j.vaccine.2018.11.005] [Medline: 30416017]
88. Zaidi SMA, Khowaja S, Kumar Dharma V, Khan AJ, Chandir S. Coverage, timeliness, and determinants of immunization completion in Pakistan: evidence from the Demographic and Health Survey (2006-07). Hum Vaccin Immunother 2014;10(6):1712-1720 [FREE Full text] [doi: 10.4161/hv.28621] [Medline: 24784118]
89. Marefiaw T, Yenesew M, Mihirete K. Age-appropriate vaccination coverage and its associated factors for pentavalent 1-3 and measles vaccine doses, in northeast Ethiopia: A community-based cross-sectional study. PLoS One 2019;14(8):e0218470 [FREE Full text] [doi: 10.1371/journal.pone.0218470] [Medline: 31419230]
90. Laryea D, Abbeyquaye Parbie E, Frimpong E. Timeliness of childhood vaccine uptake among children attending a tertiary health service facility-based immunisation clinic in Ghana. BMC Public Health 2014 Jan 29;14:90 [FREE Full text] [doi: 10.1186/1471-2458-14-90] [Medline: 24476339]
91. Cherian V, Kumar SN, Kumar SA. Assessment of timely immunization in an urbanized agglomeration of East Delhi, India. Int J Community Med Public Heal. 2019. URL: https://www.ijcmph.com/index.php/ijcmph/article/viewFile/4200/2852 [accessed 2021-06-07]
92. Menzies R, Heron L, Lampard J, McMillan M, Joseph T, Chan J, et al. A randomised controlled trial of SMS messaging and calendar reminders to improve vaccination timeliness in infants. Vaccine 2020 Mar 30;38(15):3137-3142. [doi: 10.1016/j.vaccine.2020.02.045] [Medline: 32147296]
93. Chiabi A, Nguefack F, Njapndounke F, Kobela M, Kenfack K, Nguefack S, et al. Vaccination of infants aged 0 to 11 months at the Yaounde Gynaeco-obstetric and pediatric hospital in Cameroon: how complete and how timely? BMC Pediatr 2017 Dec 19;17(1):206 [FREE Full text] [doi: 10.1186/s12887-017-0954-1] [Medline: 29258463]
94. Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 2010 Jun 14;340:c2697 [FREE Full text] [doi: 10.1136/bmj.c2697] [Medline: 20547685]
95. Shea Oscar R, Kiersten J. DHS comparative reports. The DHS Wealth Index. 2004. URL: https://dhsprogram.com/publications/publication-cr6-comparative-reports.cfm [accessed 2021-06-07]

96. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. Am J Epidemiol 1986 Jan;123(1):174-184. [doi: 10.1093/oxfordjournals.aje.a114212] [Medline: 3509965]
97. Wang B, Wang H, Tu X, Feng C. Comparisons of Superiority, Non-inferiority, and Equivalence Trials. Shanghai Arch Psychiatry 2017 Dec 25;29(6):385-388 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.217163] [Medline: 29719352]
98. 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] [Medline: 22209829]
99. Kebede A, Ajayi I, Arowojolu A. Effect of enhanced reminders on postnatal clinic attendance in Addis Ababa, Ethiopia: a cluster randomized controlled trial. Glob Health Action 2019;12(1):1609297 [FREE Full text] [doi: 10.1080/16549716.2019.1609297] [Medline: 31124401]
100. Lund S, Nielsen B, Hemed M, Boas I, Said A, Said K, et al. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. BMC Pregnancy Childbirth 2014 Jan 17;14:29 [FREE Full text] [doi: 10.1186/1471-2393-14-29] [Medline: 24438517]
101. Omole O, Ijadunola M, Olotu E, Omotoso O, Bello B, Awoniran O, et al. The effect of mobile phone short message service on maternal health in south-west Nigeria. Int J Health Plann Manage 2018 Jan;33(1):155-170. [doi: 10.1002/hpm.2404] [Medline: 28332259]
102. Stockwell M, Westhoff C, Kharbanda E, Vargas C, Camargo S, Vawdrey D, et al. Influenza vaccine text message reminders for urban, low-income pregnant women: a randomized controlled trial. Am J Public Health 2014 Feb;104 Suppl 1:e7-12. [doi: 10.2105/AJPH.2013.301620] [Medline: 24354839]
103. Matheson E, Derouin A, Gagliano M, Thompson J, Blood-Siegfried J. Increasing HPV vaccination series completion rates via text message reminders. J Pediatr Health Care 2014;28(4):e35-e39. [doi: 10.1016/j.pedhc.2013.09.001] [Medline: 24200295]
104. Hofstetter A, DuRivage N, Vargas C, Camargo S, Vawdrey D, Fisher A, et al. Text message reminders for timely routine MMR vaccination: A randomized controlled trial. Vaccine 2015 Oct 26;33(43):5741-5746 [FREE Full text] [doi: 10.1016/j.vaccine.2015.09.042] [Medline: 26424607]
105. Stockwell M, Hofstetter A, DuRivage N, Barrett A, Fernandez N, Vargas C, et al. Text message reminders for second dose of influenza vaccine: a randomized controlled trial. Pediatrics 2015 Jan;135(1):e83-e91 [FREE Full text] [doi: 10.1542/peds.2014-2475] [Medline: 25548329]

Abbreviations

EPI: expanded program on immunization
mHealth: mobile health
RR: risk ratio
VPD: vaccine-preventable disease

Edited by L Buis; submitted 30.01.21; peer-reviewed by M Das, C Seebode, F Khatun; comments to author 12.04.21; revised version received 18.04.21; accepted 11.05.21; published 15.06.21.

Please cite as:
Mekonnen ZA, Gelaye KA, Were M, Tilahun B
Effect of Mobile Phone Text Message Reminders on the Completion and Timely Receipt of Routine Childhood Vaccinations: Superiority Randomized Controlled Trial in Northwest Ethiopia
JMIR Mhealth Uhealth 2021;9(6):e27603
URL: https://mhealth.jmir.org/2021/6/e27603
doi:10.2196/27603
PMID:34128813

Original Paper

A Portable Smartphone-Based Laryngoscope System for High-Speed Vocal Cord Imaging of Patients With Throat Disorders: Instrument Validation Study

Youngkyu Kim[1,2], BSc; Jeongmin Oh[1], BSc; Seung-Ho Choi[3], MD, PhD; Ahra Jung[4], MD; June-Goo Lee[1,2], PhD; Yoon Se Lee[3], MD, PhD; Jun Ki Kim[1,2], PhD

[1]Biomedical Engineering Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea

[2]Department of Convergence Medicine, College of Medicine, University of Ulsan, Seoul, Republic of Korea

[3]Department of Otorhinolaryngology-Head and Neck Surgery, Asan Medical Center, Seoul, Republic of Korea

[4]Department of Otorhinolaryngology-Head and Neck Surgery, Eulji Medical Center, Eulji University School of Medicine, Seoul, Republic of Korea

Corresponding Author:
Jun Ki Kim, PhD
Biomedical Engineering Research Center
Asan Institute for Life Sciences
Asan Medical Center
88, Olympic-ro, 43-gil
Seoul, 05505
Republic of Korea
Phone: 82 230108619
Email: kim@amc.seoul.kr

Abstract

Background: Currently, high-speed digital imaging (HSDI), especially endoscopic HSDI, is routinely used for the diagnosis of vocal cord disorders. However, endoscopic HSDI devices are usually large and costly, which limits access to patients in underdeveloped countries and in regions with inadequate medical infrastructure. Modern smartphones have sufficient functionality to process the complex calculations that are required for processing high-resolution images and videos with a high frame rate. Recently, several attempts have been made to integrate medical endoscopes with smartphones to make them more accessible to people in underdeveloped countries.

Objective: This study aims to develop a smartphone adaptor for endoscopes, which enables smartphone-based vocal cord imaging, to demonstrate the feasibility of performing high-speed vocal cord imaging via the high-speed imaging functions of a high-performance smartphone camera, and to determine the acceptability of the smartphone-based high-speed vocal cord imaging system for clinical applications in developing countries.

Methods: A customized smartphone adaptor optical relay was designed for clinical endoscopy using selective laser melting–based 3D printing. A standard laryngoscope was attached to the smartphone adaptor to acquire high-speed vocal cord endoscopic images. Only existing basic functions of the smartphone camera were used for HSDI of the vocal cords. Extracted still frames were observed for qualitative glottal volume and shape. For image processing, segmented glottal and vocal cord areas were calculated from whole HSDI frames to characterize the amplitude of the vibrations on each side of the glottis, including the frequency, edge length, glottal areas, base cord, and lateral phase differences over the acquisition time. The device was incorporated into a preclinical videokymography diagnosis routine to compare functionality.

Results: Smartphone-based HSDI with the smartphone-endoscope adaptor could achieve 940 frames per second and a resolution of 1280 by 720 frames, which corresponds to the detection of 3 to 8 frames per vocal cycle at double the spatial resolution of existing devices. The device was used to image the vocal cords of 4 volunteers: 1 healthy individual and 3 patients with vocal cord paralysis, chronic laryngitis, or vocal cord polyps. The resultant image stacks were sufficient for most diagnostic purposes. The cost of the device including the smartphone was lower than that of existing HSDI devices. The image processing and analytics demonstrated the successful calculation of relevant diagnostic variables from the acquired images. Patients with vocal pathologies were easily differentiable in the quantitative data.

Conclusions: A smartphone-based HSDI endoscope system can function as a point-of-care clinical diagnostic device. The resulting analysis is of higher quality than that accessible by videostroboscopy and promises comparable quality and greater

accessibility than HSDI. In particular, this system is suitable for use as an accessible diagnostic tool in underdeveloped areas with inadequate medical service infrastructure.

(JMIR Mhealth Uhealth 2021;9(6):e25816) doi:10.2196/25816

KEYWORDS

smartphone; mobile phone; endoscope; high-speed imaging; vocal cord; low-cost device; mHealth; otorhinolaryngology; head and neck; throat

Introduction

The endoscope enables optical visualization of internal organs and is a fundamental diagnostic device in clinical fields; unique properties of endoscopic systems are indispensable in diagnostics [1], biopsy [2], and minimally invasive surgery [3]. To acquire more accurate and specific diagnoses, endoscope systems have been developed with higher resolution [4], greater imaging speed [5,6], quantitative image analysis [7,8], and higher performance. With these advantages, endoscopy is now considered to be an essential procedure for a variety of clinical applications, including clinical voice evaluation.

Imaging of vocal cord vibrations is one of many techniques used for clinical voice assessment. It is commonly accepted that vocal cord vibration irregularities strongly correlate with voice disorders. However, since the frequencies of vocal cord vibrations are approximately 80-240 Hz, standard videostroboscopy at 60 frames per second (fps) does not permit the capture of vocal cord movement aside from stable and periodic states [9-11]. Alternative methods have been studied to overcome the limitations of videostroboscopy for imaging vocal cord vibrations. The most promising approach is high-speed digital imaging (HSDI) [12], which typically captures images between 4000 and 8000 fps [13]. The high frame rate of HSDI allows for the capture of vocal vibrations from 80-240 Hz, which cannot be observed in standard 60 fps imaging systems. Furthermore, HSDI provides a sufficient resolution of more than 256 × 256 pixels [14]. Although the resolution is hardly comparable to that of HD (high definition), it is sufficient for medical analysis. Use of HSDI for clinical voice evaluation has been extensively demonstrated [10,15,16]. However, despite several advantages, HSDI is not widely used in clinical diagnosis because systems with better performance than laryngostroboscopy are generally expensive, require specialized technology, and have large data footprints [13].

Endoscopic diagnosis is mostly inaccessible for patients in underdeveloped and developing countries. Indeed, even in the most developed and well-equipped hospitals in Nigeria, there may be only one functional gastroscope and/or colonoscope, for which the accessories are often deficient. Furthermore, the majority of the teaching hospitals in Nigeria have no facilities for therapeutic endoscopy [17]; this leads to a lack of experience in endoscopy for medical students, which can lead to unsatisfactory or even fatal endoscopy procedures. Furthermore, more serious issues, including inadequate maintenance, cross-infections, and insufficient performance of devices, have also been found in developing countries. In facilities with insufficient endoscopic gastroscope and colonoscope accessories, HSDI devices are difficult to furnish because of their high cost and additional technology requirements. This seriously restricts clinical voice evaluation in underdeveloped countries since HSDI has become an important part in the diagnosis of voice disorders. As with many problems regarding endoscopy in developing countries, affordable and more manageable endoscopic HSDI systems are urgently required.

Several point-of-care diagnostic devices have been invented in response to the demand in developing and underdeveloped countries. Prior to the age of integrated electronics, handheld point-of-care devices were largely restricted to diagnostic strips, which relied heavily on color changes due to chemical reactions [18]. With the ubiquity of smartphones, the rate of invention of electronic point-of-care devices has drastically increased. Modern smartphones offer a computing performance comparable to that of personal computers, and imaging quality similar to that of professional camera devices. Consequently, there have been many attempts to use smartphones for point-of-care diagnosis. For example, smartphone cameras can be used to distinguish reaction intensities, and several smartphone attachments have been introduced for biological analysis similar to that performed by enzyme-linked immunosorbent assay (ELISA) [19], fecal hemoglobin detection [20], and electrochemical monitoring of blood contents [21]. This represents an evolution of traditional point-of-care strip diagnosis, in which a smartphone can be used to distinguish the intensity of a biological reaction. Other proposed applications of the smartphone camera for point-of-care diagnosis include clinical microscopy, endoscopic imaging, video nasolaryngoscopy, flexible robotic endoscopy, etc [22-28].

In this study, we introduce a smartphone-based endoscopic imaging device and validate its imaging performance relative to videostroboscopy and HSDI. The smartphone adaptor was designed and manufactured using 3D design software and selective laser melting (SLM)–based 3D printing. This customized smartphone adaptor for the endoscope has minimal and low-cost optical components for easy maintenance and control. Unlike previous studies that only used smartphones for acquisition of regular images and videos, in this study, the specialized high-speed imaging function of the smartphone was applied for HSDI of vocal cord vibrations. With this functionality, high-speed vocal cord videos were acquired at 940 fps following standard clinical protocols and analyzed using postprocessed imaging techniques such as segmentation and registration to simplify diagnosis. The resultant images are quantitatively more revealing than videostroboscopy, and approach the quality of HSDI due to their high resolution and frame rate. In preclinical image analysis, several diagnostic variables were determined likely to assist in the diagnosis of common vocal cord pathologies. This custom simple, low-cost

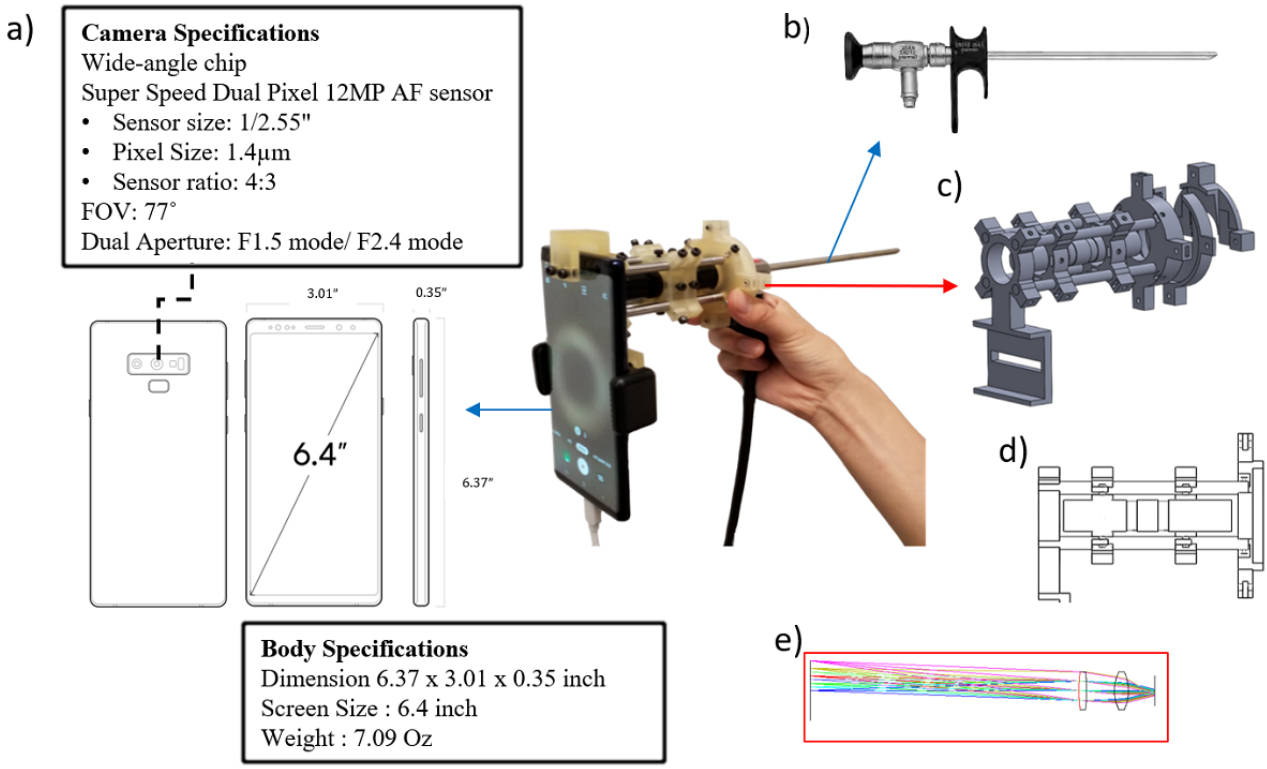
adaptor promises to reduce the cost of high-speed clinical imaging by an order of magnitude when it can be used as a substitute for standard HSDI. The device demonstrates the promise and emerging capabilities of clinical diagnostic tools that incorporate commodity smartphone technology to bring point-of-care diagnosis to remote locations.

Methods

Development of a Smartphone-Adapted Endoscope

A smartphone adaptor was developed for use in clinical endoscopy (Figure 1) with a standard 70° rigid laryngoscope (Karl Storz Co). The adaptor consists of a smartphone holder, lenses, a lens tube, and mounts for the lens tube and for connection to the endoscope. The holders and connectors were designed using 3D modeling software (Solid Works) and printed by an SLM 3D printer (Objet260, Stratasys Ltd). A lens system was used to magnify the endoscope probe eyepiece image between the endoscope eyepiece and the smartphone camera and was held by 3D-printed parts. To make the system more cost-effective, a combination of 2 biconvex lenses was used instead of expensive achromatic lenses for magnification of the endoscopic image. The focal lengths of the 2 biconvex lenses were 50 mm and 15 mm, respectively; this optical setup could acquire images with approximately 1.5× magnification. Illumination was coupled into the illumination port on the endoscope. For endoscopic illumination, we used a broadband (360-770 nm) LED light source (X-Cite XYLIS, Excelitas Technologies Corp), which was connected to the illumination port of the endoscope with a liquid light guide.

Figure 1. Schematics of a customized smartphone-endoscope adaptor that enables high-speed laryngoscopy. (A) Smartphone camera and body specifications. (B) A photograph of the demonstrated smartphone adaptor incorporating a rigid clinical endoscope. (C) 3D models of the customized smartphone-endoscope adaptor. (D) Cross-sectional view of the customized smartphone-endoscope adaptor. (E) Schematic and lens simulations of the magnification lens system. AF: autofocus, FOV: field of view.



A specific smartphone model (Galaxy Note 9, Samsung) was used to capture endoscopic images at a high frame rate. The catalog specification provided by the manufacturer of the smartphone claims that this specific smartphone model can acquire high-speed images at 960 fps, for a maximum duration of 0.4 seconds, at 1280 × 720 pixel resolution [29]. The complementary metal-oxide-semiconductor sensor of the smartphone is specified for 12.0 megapixels (MP) and 1/2.55 inches (1.4 µm pixel size), and is packaged with a dual aperture mode.

Statistical and Data Analysis

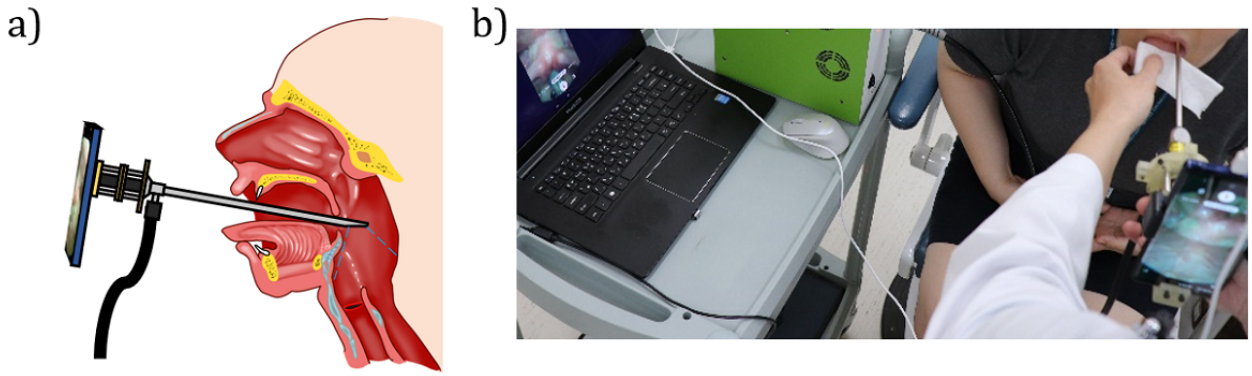
The acquired high-speed vocal cord vibration images were segmented using the seeded region growing algorithm [30] to determine the amplitude of the vibrations on each side of the glottis [31]. Image analysis software was developed in MATLAB (MathWorks) to calculate the diagnostic parameters, including the vocal cord vibration frequencies, glottal edge phase shifts, and total glottal area changes.

Clinical Experiment

The developed smartphone-based imaging system was used to acquire a high-speed video of the vocal cords of 4 human subjects—1 healthy individual and 3 patients with known vocal pathologies. A board-certified otolaryngologist at Asan Medical Center with 15 years of experience followed common videokymography routines [11] to capture high-speed diagnostic videos in a clinical environment (Figure 2). The routine vocal cord examination was as follows: the first step was to interview the patient and assess their symptoms and voice status, and the

second step was videostrobolaryngoscopy or videokymography with an endoscope. Depending on the patient's condition, additional tests such as computed tomography (CT) scans or a blood test may be required.

Figure 2. (A) A schematic representation of laryngeal imaging performed by the smartphone-based high-speed imaging system. (B) A clinical demonstration of the smartphone-based high-speed imaging system.



Before imaging human subjects, the smartphone-based HSDI system was adapted to the clinical endoscope by adjusting the fine focus and fixing the endoscope in place. Following confirmation that the received images were clear and focused, the clinician performed routine vocal cord evaluation using the endoscope-smartphone device. When the diagnosis routine was completed, the endoscope probe was wiped with a disinfectant to avoid cross-contamination between human subjects [32]. The whole procedure took less than a minute, thus avoiding undue burden on volunteers.

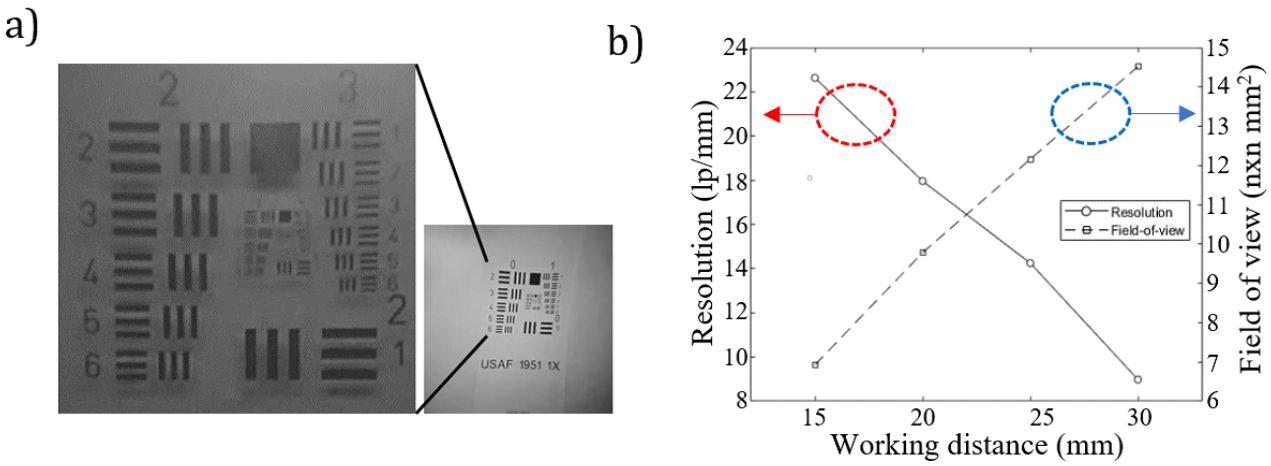
The study was approved by the Institutional Review Board (IRB #2020-0798) of Asan Medical Center, Seoul, under the Korean Bioethics and Safety Act and the Korean Medical Device Act.

Results

System Characterization

The US Air Force resolution target was used to evaluate the imaging performance of the smartphone-endoscope system (Figure 3A). The distance between the end of the endoscope probe and the resolution target was adjusted from 15 mm to 30 mm in order to determine the proper optical field of view (FOV) and resolution. At the highest resolution, the system resolved 23 line pairs per millimeter (lp/mm), which corresponded to 43.478 μm at a working distance of 15 mm. The resolution of the system was observed to decrease as the working distance was increased. In addition, the optical FOV was 7.8 mm × 7.8 mm at a 15 mm working distance, and 14.6 mm × 14.6 mm at a 30 mm working distance; thus the optical FOV increased approximately linearly with the working distance (Figure 3B).

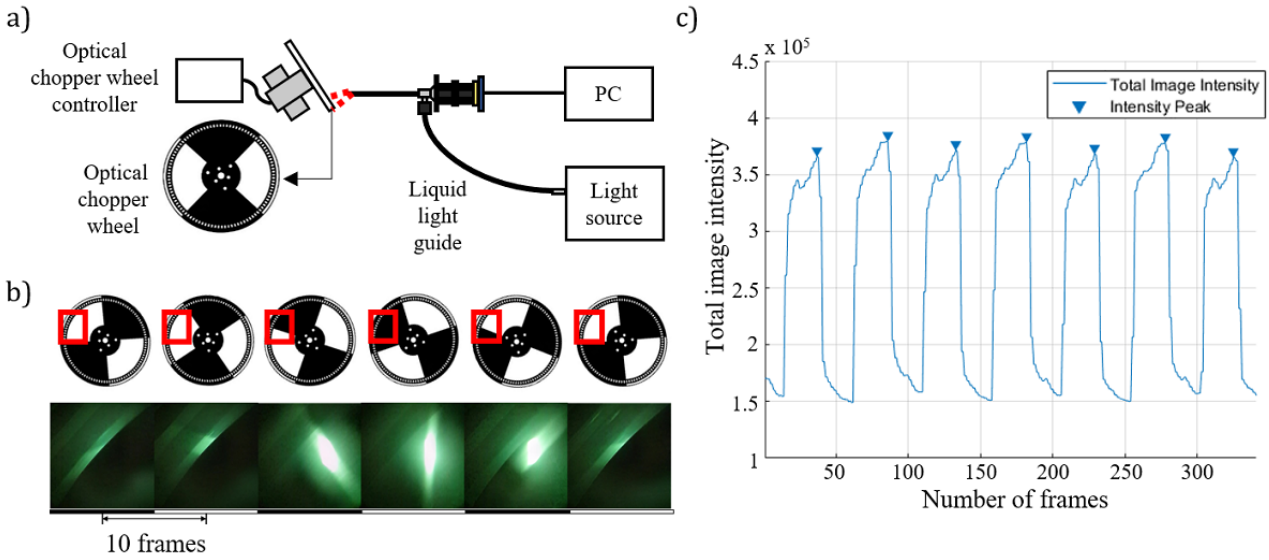
Figure 3. (A) US Air Force target test image taken at 960 fps. (B) A graph of the measured image resolution and field of view as functions of working distance.



A custom frame rate test was designed to confirm the video frame rate of the customized high-speed system. An optical chopper wheel was used to create a time-dependent moving sample of known frequency. Two quadrants of a 4-quadrant optical chopper wheel mask were optically blocked, and the other 2 quadrants were left transparent as shown schematically in Figure 4. While the light source provided illumination from the endoscope, high frame rate videos were acquired with the optical chopper wheel rotating at 10 Hz. The acquired high frame rate videos were analyzed by image processing in

MATLAB, as light reflected from the chopper wheel was captured by the high-speed imaging system. The total intensity of light reflected from the 2 quadrants of the chopper wheel resulted in a series of image frames with a periodically modulated intensity signal that could be plotted frame by frame. The period of intensity modulation was calculated as the time interval between frames with peak intensities and was equal to half the rotational period of the optical chopper; this was measured to be 46.972 frames. Since the optical chopper was rotating at 10 Hz, the duration of a half-rotation was 0.05 seconds; therefore, the actual frame rate of the smartphone was 939.44 fps.

Figure 4. (A) Schematic of the system for high-speed imaging assessment. (B) Image series acquired over 50 frames. (C) A plot of the total image intensity by chopper wheel rotation. PC: personal computer.



Clinical Vocal Cord Vibration HSDI

The obtained HSDI vocal cord vibration videos were taken through a clinical diagnosis process. The results are presented in Figures 5 and 6. As depicted in Figure 5, the 940-fps HSDI video obtained from the clinical experiments was segmented into vocal cord area and glottal area for further analysis; this allowed us to obtain the parameters of the vocal cord with the largest contribution to clinical diagnosis. The parameters obtained from the vocal cord vibration HSDI data included the base vibration frequency of the vocal cord, the difference between the left glottal area and the right glottal area, and the total glottal area.

As shown in Figure 5, the data were obtained from the healthy subject, and patients with left vocal cord paralysis, chronic laryngitis, and right vocal cord polyp. The original RGB (red-green-blue) color image data were converted to grayscale, and the contrast was adjusted for glottal area segmentation using MATLAB. The seeded region growing method [30] was used to segment the glottal area from the image processed data. In addition, the anatomical midline of the vocal cord was drawn, after consultation with a clinician, to divide the left and right glottal areas.

The segmented glottal areas were analyzed in order to determine any differences between the data of the healthy subject and those of patients. From healthy subject data, normalized total glottal area changes are shown in Figure 6A. Using the fast Fourier transform, we were able to remove motion error and extract the base frequency of the healthy vocal cord. The fundamental frequency of the healthy subject, a 24-year-old female, was 224 Hz. The average fundamental frequency of adult females of this age is reported to be similar, at around 217 Hz [33]. Figure 6B shows the normalized total and lateral (left and right) glottal area changes in the vocal cords of a patient with left vocal cord paralysis. These data show that the amplitudes of the left vocal cord vibrations are significantly lower than those of the right vocal cord. The normalized total and lateral glottal areas of the patient with chronic laryngitis are shown in Figure 6C. Although the maximum amplitudes are similar to those of the healthy subject in both vocal cords, the minimum amplitude of each cycle differs from those of the healthy subject. The minimum amplitude value of each cycle in the healthy subject had almost zero area; this is because the vocal cord closes completely when the vocalizations are generated in a healthy subject. However, in the patient with chronic laryngitis, the minimum glottal area was larger than that in the healthy subject, because in chronic laryngitis the vocal cords do not close completely as sound is generated. Figure 6D shows the normalized total and lateral glottal areas for the vocal cords of a patient with a polyp on the right vocal cord. In this case, the vibrations of the right vocal cord were restricted by the polyp; therefore, the vibration amplitude of the right glottal area is less than that of the left glottal area. Similar to chronic laryngitis, the patient with the right vocal cord polyp had much larger minimum areas than the healthy subject; this phenomenon is caused by imperfect closure of the vocal cords due to the polyp.

Figure 5. High-speed images of the glottis and their segmentation. Raw vocal cord image (left), preprocessed image for segmentation (center), and area of the normal vocal cord by image segmentation (right). (A) The glottis of a healthy volunteer. (B) Images taken from a patient with left vocal cord paralysis. (C) Vocal analysis from a patient with chronic laryngitis. (D) Images from a patient with a right vocal cord polyp.

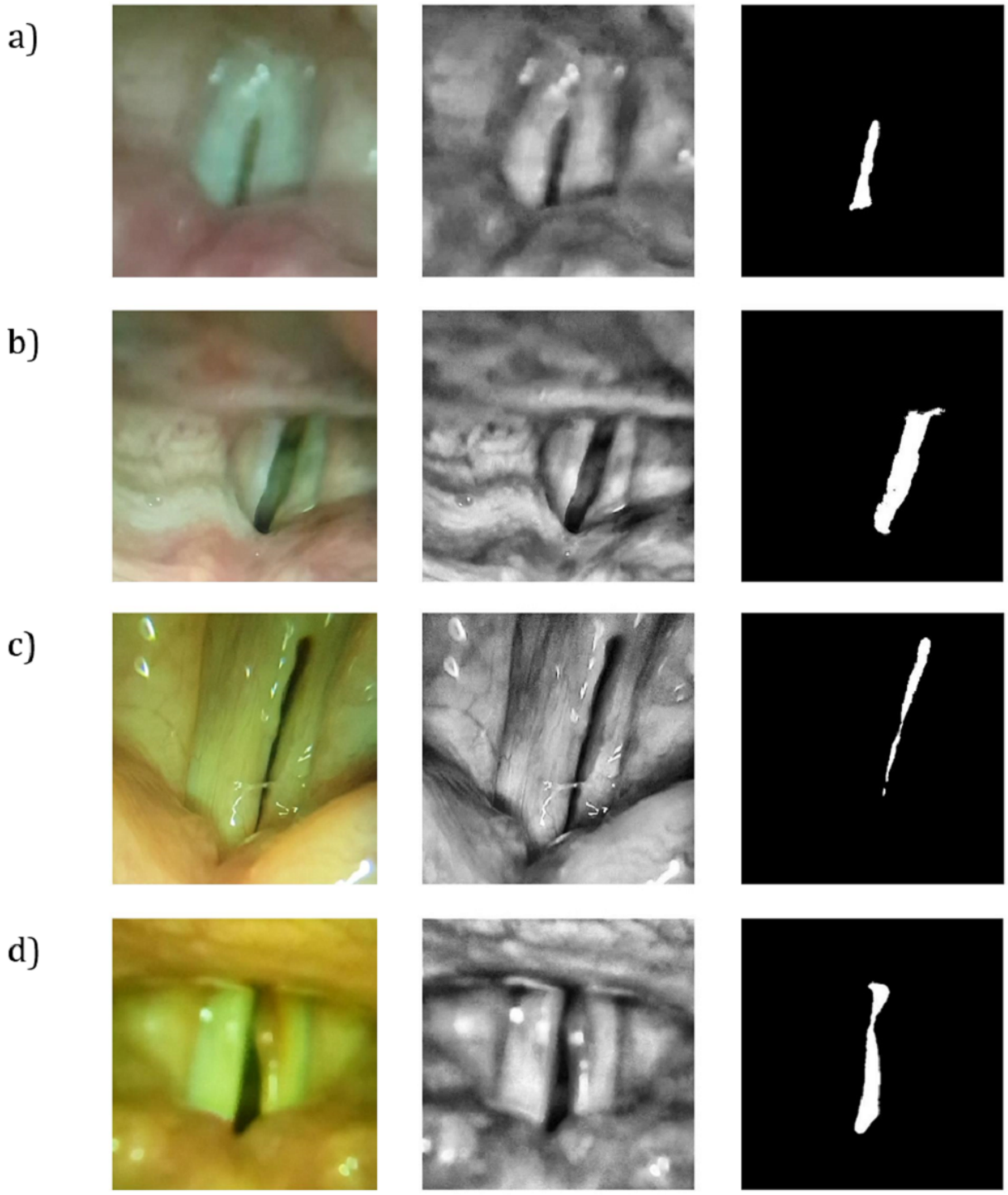
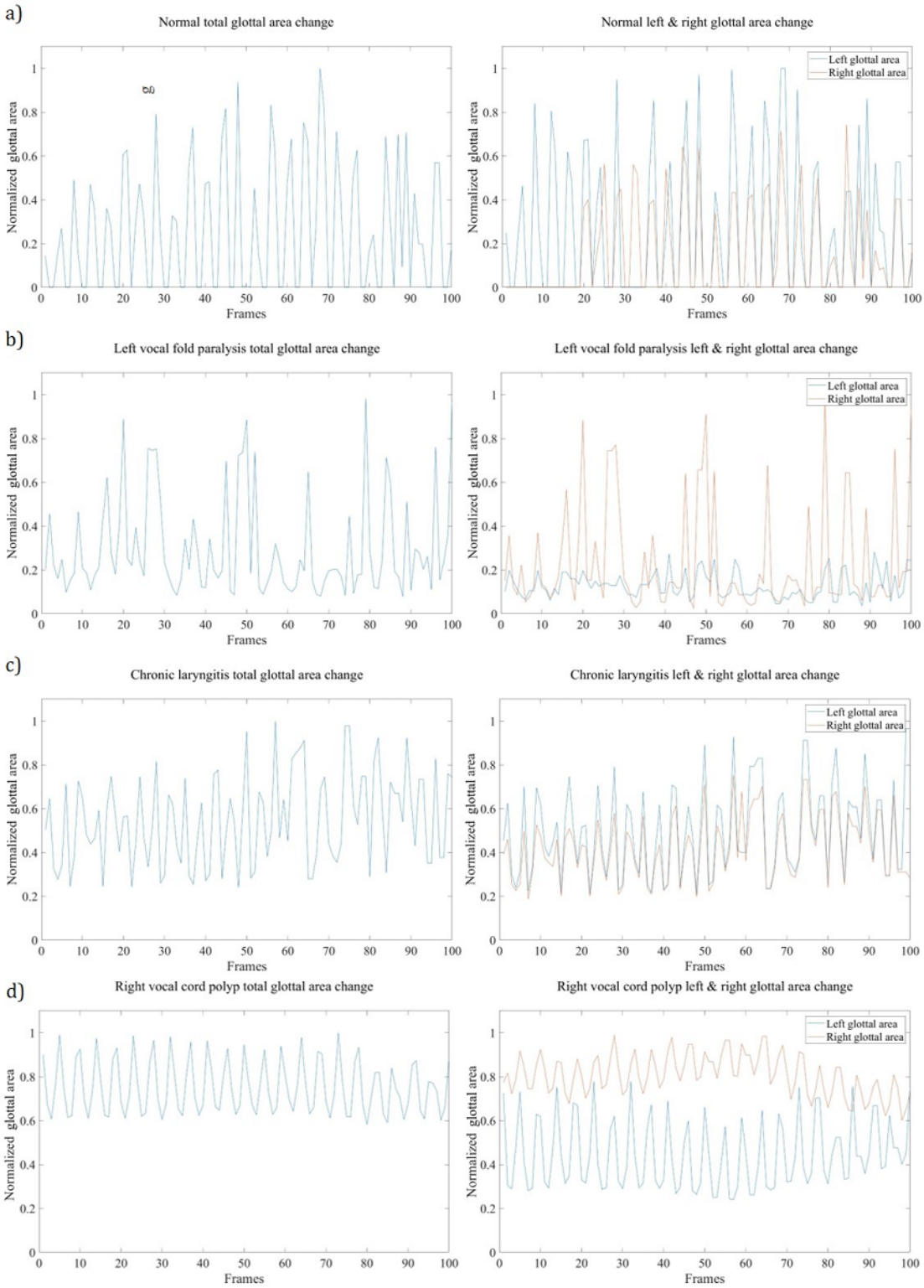


Figure 6. Plots of the normalized total glottal area (left) and a comparison of the normalized left and right glottal areas separated by the anatomical midline of the glottis over the course of a 0.4-second video at 940 fps (right) plotted for each patient: (A) normal healthy subject, (B) patient with left vocal cord paralysis, (C) patient with chronic laryngitis, and (D) patient with right vocal cord polyp.



Discussion

Principal Findings

Although many point-of-care devices have been developed based on smartphone systems, the full use of smartphone capabilities has been limited. Most smartphone-based point-of-care systems use chemical reactions of bodily fluids, including blood, saliva, sweat, and urine [18,34-36], and usage of smartphone functionalities are restricted to recognition of color changes or receiving electric signals generated by biochemical reactions. Most point-of-care devices are invasive, require bodily fluids, and their methods have limited applicability for noninvasive diagnostics. In these cases, the function of the camera is restricted to basic image acquisition

[34]. Another way to integrate the smartphone into a point-of-care device would be by using the built-in functionality of a smartphone. Recently, consumer demand has resulted in rapid upgrades to embedded smartphone camera technology, resulting in unique functionality that can be incorporated into a special purpose imaging device.

Although commercialized HSDI devices enable unparalleled diagnoses, they also have several limitations. Their high cost and need for additional connection to video transmission processors like PCI-e (Peripheral Component Interconnect Express) frame grabbers make their application difficult for users in underdeveloped countries. In comparison, our smartphone-based HSDI system combines a printed endoscope-smartphone adapter and a low-cost commercial lens, demonstrating that even with low-cost components, it is possible to acquire high-speed digital images of vocal cord vibrations with existing clinical endoscopes.

Normally, images of vocal cord vibrations obtained by HSDI are postprocessed before being used in actual clinical practice. First, masking areas are drawn manually by the clinician, and the area is then plotted or processed for clinically relevant parameters. While this is traditionally performed by examining a series of individual frames superimposed on a laryngeal paralysis stroboscopic recording, or digital kymograms superimposed on wavefronts [37-39], modern computing makes it possible to extract glottal areas frame by frame using semiautomated image segmentation. As a result, the vocal cord areas and the glottal areas are divided and segmented as shown in Figure 5A. The clinical diagnostics obtained by the above process are the base cord and lateral phase difference of the vocal cord. The base cord is defined as the amount of change in the area of the entire glottal region, while the lateral phase difference is defined as the amount of change in the edge length of the left and right glottal areas [40].

The patient data obtained from our device made it possible to identify some well-known differences between patients with voice pathologies and the healthy subject. These results suggest that the smartphone-based HSDI equipment has the potential to be used to diagnose various voice diseases.

We first obtained clinical control data from a healthy female volunteer. The fundamental frequency appeared to be 224 Hz, which is similar to the average vocal cord fundamental frequency of adult females in this age range, which has been reported to be 217 Hz [33]. Although typical HSDI guidelines call for 7 to 10 frames per vibrational period, in this demonstration vocal cord period, amplitudes at normal speaking frequencies were successfully captured with 3- to 4-fold frame rates, demonstrating the suitability of this technology for clinical diagnostics.

Several clinically relevant features that differentiated the healthy control from the patients with voice pathologies were successfully observed. As shown in Figure 6, diagnosis of glottal pathologies can be easily obtained from plots of the vocal cord areas as calculated by segmentation of the glottal area in the images. In all patients, vocal impairments were shown by the limited amplitude of glottal vibration. In the patient with the one-sided vocal cord paralysis, the diagnosis presented as severely limited vibrational amplitude on one side, while chronic laryngitis presented with limited closure of the vocal cords. A polyp on one side of the focal cord displayed both limited closure and limited vibrational amplitude. Although our system does not achieve the high frame rate of clinical videokymography [37], the data analysis achieved from a 0.4 second clip is comparable in quality to that of videostroboscopy [38,41], capturing phase shifts of vibration, amplitude, vibrational asymmetry, and characteristic immobilization for diagnosis.

Commercially available HSDI systems, as shown in Table 1, have list prices of more than $10,000, and require installation of PCI cards to acquire an image series. Conversely, the smartphone-based HSDI system developed in this study can be configured for less than $500, which includes manufacturing of all 3D-printed parts and lens optics. The manufacturing costs of plastic parts could be further reduced by mass production. Furthermore, no additional connection interfaces are required to capture video, since the video interface used is included in the smartphone software. These calculations exclude the price of the smartphone, as smartphones are widely distributed in underserved markets. However, even if specific models of smartphone are required for smartphone-based HSDI systems, the price of a new smartphone is significantly less than that of a standard HSDI system. In addition, since smartphones can be connected to both Wi-Fi and mobile internet, users can easily transfer the captured video via wireless communication for analysis in remote regions. In developing countries, large hospitals may be significantly removed from the point of care, and this mobility can help to distribute medical service more equitably. Our system is simple and has intuitive maintenance requirements; therefore, it is easier to educate medical students in countries where medical schools are limited, although requirements for the cleaning of endoscopes complicate training and distribution.

Table 1. Comparison between a commercial high-speed digital imaging (HSDI) system and the smartphone-based HSDI system.

Characteristic	Commercial HSDI system (FASTCAM MC2, Photron)	Smartphone-based HSDI system
Frame rate	• 4000 fps	• 940 fps
Pixel count	• 512 × 512	• 1280 × 720
Extra connection interface	• Gigabit Ethernet	• USB 3.0 port
Price	• >USD $10,000, excluding computer	• <USD $400, excluding smartphone
Size	• Camera: 35 mm × 35 mm × 34 mm • Processor: 195 mm × 159 mm × 130 mm	• Smartphone: 76 mm × 162 mm × 9 mm • Adaptor: 115 mm × 120 mm × 80 mm
Weight	• Camera: 100 g • Processor: 5000 g • Total : 5100 g	• Smartphone: 201 g • Adaptor: 295 g • Total: 496 g

In addition, the smartphone-based HSDI system has a higher resolution than a traditional videostroboscope. Although the main focus of HSDI is a high frame rate, higher resolution enables the user to observe details that can assist with diagnosis. However, the advantages of mobility and price come with trade-offs in the form of device performance. Due to the limitations of the built-in smartphone functions, images can only be acquired at a maximum of 940 fps, which is considerably lower than that of the commercial HSDI system, which can achieve 4000 to 8000 fps. Typically, the fundamental vocal cord frequency for patients aged 17 to 25 years ranges from 115-132 Hz for males and 200-260 Hz for females [33]. Therefore, the smartphone-based system only captures 3 to 8 frames per vibration cycle, whereas 15 to 34 frames would be captured per vibration with the traditional HSDI. This frame rate is sufficiently high to detect the tendency of changes in the vocal cord area, but not sufficient to observe the details of each vibration cycle. This restriction may omit the detail necessary to diagnose certain conditions of the vocal cord. Therefore, additional filtering is required for frequency analysis.

Furthermore, the limitation on image capture duration to only 0.4 seconds at a time results in some pitfalls for clinical imaging, as it may be necessary for clinicians to repeat imaging in order to secure sufficient data for clinical evaluation. Although 0.4 seconds corresponds to 376 frames (47 vibrations), a clinician may have to take several HSDI videos from the smartphone to obtain sufficient data for analysis. Another drawback of the smartphone-based HSDI system developed in this study is that it requires a stronger endoscopic light source. As the charge-coupled device of smartphone cameras has lower photosensitivity than commercially available HSDI systems, a normal endoscope light source would not yield quality data. The stronger light source requires more power and can cause heat dissipation to become a consideration in the design of the device.

In the future, to overcome the aforementioned limitations, the development of more compact and optically shielded adaptors will be required. Emerging imaging solutions such as artificial intelligence–based image processing and deep learning could be applied to HSDI images of the vocal cord to overcome inadequate frame rates and acquisition times.

Conclusions

Recent advances in smartphone technology, particularly in the domain of image sensors, have produced technology which rivals the performance of existing commercial high-speed cameras. In this study, we investigated whether the HSDI system currently used for vocal cord assessment can be replaced to some extent by using the high-speed video acquisition function from a smartphone camera. In validating the imaging capabilities of our smartphone-based high-speed endoscopy system against existing HSDI technology, our results demonstrate that it is possible to process clinical images for useful diagnostic reference values of healthy subjects. Furthermore, the possibility of disease diagnosis for actual patients was presented. Further development of smartphone-based HSDI technology is necessary to improve the distribution of health care capabilities.

Acknowledgments

This work was supported by the National Research Foundation of Korea (2019R1A2C2084122) and MRC grant (2018R1A5A2020732) by the Ministry of Science & ICT, as well as by the Ministry of Trade, Industry, and Energy under the Industrial Technology Innovation Program (10080726, 20000843) and by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (HI18C2391).

Conflicts of Interest

None declared.

References

1. Rosen CA, Murry T. Diagnostic laryngeal endoscopy. Otolaryngologic Clinics of North America 2000 Aug;33(4):751-757. [doi: 10.1016/s0030-6665(05)70241-3]
2. van Rossum PS, Goense L, Meziani J, Reitsma JB, Siersema PD, Vleggaar FP, et al. Endoscopic biopsy and EUS for the detection of pathologic complete response after neoadjuvant chemoradiotherapy in esophageal cancer: a systematic review and meta-analysis. Gastrointest Endosc 2016 May;83(5):866-879. [doi: 10.1016/j.gie.2015.11.026] [Medline: 26632523]
3. Davila RE, Rajan E, Adler D, Hirota WK, Jacobson BC, Leighton JA, et al. ASGE guideline: the role of endoscopy in the diagnosis, staging, and management of colorectal cancer. Gastrointestinal Endoscopy 2005 Jan;61(1):1-7. [doi: 10.1016/s0016-5107(04)02391-0]
4. Schlegel P, Kunduk M, Stingl M, Semmler M, Döllinger M, Bohr C, et al. Influence of spatial camera resolution in high-speed videoendoscopy on laryngeal parameters. PLoS One 2019 Apr 22;14(4):e0215168 [FREE Full text] [doi: 10.1371/journal.pone.0215168] [Medline: 31009488]
5. Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. Clinical implementation of laryngeal high-speed videoendoscopy: challenges and evolution. Folia Phoniatr Logop 2008 Nov 30;60(1):33-44 [FREE Full text] [doi: 10.1159/000111802] [Medline: 18057909]
6. Poburka BJ, Patel RR, Bless DM. Voice-Vibratory Assessment With Laryngeal Imaging (VALI) Form: Reliability of Rating Stroboscopy and High-speed Videoendoscopy. J Voice 2017 Jul;31(4):513.e1-513.e14. [doi: 10.1016/j.jvoice.2016.12.003] [Medline: 28040342]
7. Mehta DD, Deliyski DD, Quatieri TF, Hillman RE. Automated Measurement of Vocal Fold Vibratory Asymmetry From High-Speed Videoendoscopy Recordings. J Speech Lang Hear Res 2011 Feb;54(1):47-54. [doi: 10.1044/1092-4388(2010/10-0026)]
8. Bohr C, Kraeck A, Eysholdt U, Ziethe A, Döllinger M. Quantitative analysis of organic vocal fold pathologies in females by high-speed endoscopy. Laryngoscope 2013 Jul 06;123(7):1686-1693. [doi: 10.1002/lary.23783] [Medline: 23649746]
9. Bless D, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. Ear Nose Throat J 1987 Jul;66(7):289-296. [Medline: 3622324]
10. Hirose H. High-speed digital imaging of vocal fold vibration. Acta Otolaryngol Suppl 1988 Jul 08;458(sup458):151-153. [doi: 10.3109/00016488809125120] [Medline: 3245422]
11. Schutte HK, Svec JG, Sram F. First results of clinical application of videokymography. Laryngoscope 1998 Aug;108(8 Pt 1):1206-1210. [doi: 10.1097/00005537-199808000-00020] [Medline: 9707245]
12. Schade G, Müller F. [High speed glottographic diagnostics in laryngology]. HNO 2005 Dec;53(12):1085-6, 1088. [doi: 10.1007/s00106-005-1285-3] [Medline: 15951995]
13. Echternach M. High-Speed Digital Imaging. In: Sataloff RT, editor. Neurolaryngology. San Diego, CA: Plural Publishing; 2017:369.
14. Guzman M, Laukkanen A, Traser L, Geneid A, Richter B, Muñoz D, et al. The influence of water resistance therapy on vocal fold vibration: a high-speed digital imaging study. Logoped Phoniatr Vocol 2017 Oct 02;42(3):99-107. [doi: 10.1080/14015439.2016.1207097] [Medline: 27484690]
15. Lohscheller J, Dollinger M, Schuster M, Schwarz R, Eysholdt U, Hoppe U. Quantitative Investigation of the Vibration Pattern of the Substitute Voice Generator. IEEE Trans Biomed Eng 2004 Aug;51(8):1394-1400. [doi: 10.1109/tbme.2004.827938]
16. Yamauchi A, Yokonishi H, Imagawa H, Sakakibara K, Nito T, Tayama N, et al. Quantification of Vocal Fold Vibration in Various Laryngeal Disorders Using High-Speed Digital Imaging. J Voice 2016 Mar;30(2):205-214. [doi: 10.1016/j.jvoice.2015.04.016] [Medline: 26003886]
17. Nwokediuko SC. Challenges of Gastrointestinal Endoscopy in Resource-Poor Countries. In: Gastrointestinal Endoscopy. Norderstedt, Deutschland: Books on Demand; Nov 21, 2011.
18. St John A, Price C. Existing and Emerging Technologies for Point-of-Care Testing. Clin Biochem Rev 2014 Aug;35(3):155-167 [FREE Full text] [Medline: 25336761]
19. del Rosario M, Redmond S, Lovell N. Tracking the Evolution of Smartphone Sensing for Monitoring Human Movement. Sensors (Basel) 2015 Jul 31;15(8):18901-18933 [FREE Full text] [doi: 10.3390/s150818901] [Medline: 26263998]
20. Soraya G, Nguyen T, Abeyrathne C, Huynh D, Chan J, Nguyen P, et al. A Label-Free, Quantitative Fecal Hemoglobin Detection Platform for Colorectal Cancer Screening. Biosensors (Basel) 2017 May 05;7(2):19 [FREE Full text] [doi: 10.3390/bios7020019] [Medline: 28475117]
21. Guo J. Smartphone-Powered Electrochemical Dongle for Point-of-Care Monitoring of Blood β-Ketone. Anal Chem 2017 Sep 05;89(17):8609-8613. [doi: 10.1021/acs.analchem.7b02531] [Medline: 28825471]
22. Breslauer DN, Maamari RN, Switz NA, Lam WA, Fletcher DA. Mobile phone based clinical microscopy for global health applications. PLoS One 2009 Jul 22;4(7):e6320 [FREE Full text] [doi: 10.1371/journal.pone.0006320] [Medline: 19623251]
23. Mistry N, Coulson C, George A. endoscope-i: an innovation in mobile endoscopic technology transforming the delivery of patient care in otolaryngology. Expert Rev Med Devices 2017 Nov 17;14(11):913-918. [doi: 10.1080/17434440.2017.1386548] [Medline: 28972409]

24. Bae JK, Vavilin A, You JS, Kim H, Ryu SY, Jang JH, et al. Smartphone-Based Endoscope System for Advanced Point-of-Care Diagnostics: Feasibility Study. JMIR Mhealth Uhealth 2017 Jul 27;5(7):e99 [FREE Full text] [doi: 10.2196/mhealth.7232] [Medline: 28751302]
25. Ha JHI, Sagili SR. Smartphone adaptor use for nasal endoscopy. Eye (Lond) 2019 May 16;33(5):854-855 [FREE Full text] [doi: 10.1038/s41433-018-0334-6] [Medline: 30651594]
26. Quimby AE, Kohlert S, Caulley L, Bromwich M. Smartphone adapters for flexible Nasolaryngoscopy: a systematic review. J Otolaryngol Head Neck Surg 2018 May 08;47(1):30-37 [FREE Full text] [doi: 10.1186/s40463-018-0279-6] [Medline: 29739440]
27. Moon Y, Oh J, Hyun J, Kim Y, Choi J, Namgoong J, et al. Cost-Effective Smartphone-Based Articulable Endoscope Systems for Developing Countries: Instrument Validation Study. JMIR Mhealth Uhealth 2020 Sep 10;8(9):e17057 [FREE Full text] [doi: 10.2196/17057] [Medline: 32909951]
28. Agu E, Pedersen P, Strong D, Tulu B, He Q, Wang L, et al. The smartphone as a medical device: Assessing enablers, benefits and challenges. 2013 Presented at: IEEE International Workshop of Internet-of-Things Networking and Control (IoT-NC); June 24; New Orleans, LA p. 48-52. [doi: 10.1109/IoT-NC.2013.6694053]
29. The Official Samsung Galaxy Site: Specifications. Samsung. 2018. URL: https://www.samsung.com/global/galaxy/galaxy-note9/specs/ [accessed 2020-09-11]
30. Adams R, Bischof L. Seeded region growing. IEEE Trans. Pattern Anal. Machine Intell 1994 Jun;16(6):641-647. [doi: 10.1109/34.295913]
31. Schlegel P, Kniesburges S, Dürr S, Schützenberger A, Döllinger M. Machine learning based identification of relevant parameters for functional voice disorders derived from endoscopic high-speed recordings. Sci Rep 2020 Jun 29;10(1):10517 [FREE Full text] [doi: 10.1038/s41598-020-66405-y] [Medline: 32601277]
32. Moses FM, Lee JS. Current GI endoscope disinfection and QA practices. Dig Dis Sci 2004 Nov 1;49(11-12):1791-1797. [doi: 10.1007/s10620-004-9572-5] [Medline: 15628705]
33. Huss PJ. Vocal Pitch Range and Habitual Pitch Level: The Study of Normal College Age Speakers. Master's Theses.Western Michigan University 1983:1590 [FREE Full text]
34. Xu X, Akay A, Wei H, Wang S, Pingguan-Murphy B, Erlandsson B, et al. Advances in Smartphone-Based Point-of-Care Diagnostics. Proc. IEEE 2015 Feb;103(2):236-247. [doi: 10.1109/JPROC.2014.2378776]
35. Roda A, Michelini E, Zangheri M, Di Fusco M, Calabria D, Simoni P. Smartphone-based biosensors: A critical review and perspectives. TrAC Trends in Analytical Chemistry 2016 May;79:317-325. [doi: 10.1016/j.trac.2015.10.019] [Medline: 25904163]
36. Liu J, Geng Z, Fan Z, Liu J, Chen H. Point-of-care testing based on smartphone: The current state-of-the-art (2017-2018). Biosens Bioelectron 2019 May 01;132:17-37. [doi: 10.1016/j.bios.2019.01.068] [Medline: 30851493]
37. Eysholdt U, Tigges M, Wittenberg T, Pröschel U. Direct evaluation of high-speed recordings of vocal fold vibrations. Folia Phoniatr Logop 1996;48(4):163-170. [doi: 10.1159/000266404] [Medline: 8823984]
38. Sercarz JA, Berke GS, Ming Y, Gerratt BR, Natividad M. Videostroboscopy of human vocal fold paralysis. Ann Otol Rhinol Laryngol 1992 Jul 28;101(7):567-577. [doi: 10.1177/000348949210100705] [Medline: 1626902]
39. Wittenberg T, Tigges M, Mergell P, Eysholdt U. Functional imaging of vocal fold vibration: Digital multislice high-speed kymography. Journal of Voice 2000 Sep;14(3):422-442. [doi: 10.1016/s0892-1997(00)80087-9]
40. Sulter AM, Schutte HK, Miller DG. Standardized laryngeal videostroboscopic rating: Differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. Journal of Voice 1996 Jan;10(2):175-189. [doi: 10.1016/s0892-1997(96)80045-2]
41. Banjara H, Mungutwar V, Singh D, Gupta A, Singh S. Demographic and videostroboscopic assessment of vocal pathologies. Indian J Otolaryngol Head Neck Surg 2012 Jun 6;64(2):150-157 [FREE Full text] [doi: 10.1007/s12070-011-0451-z] [Medline: 23730576]

Abbreviations

CT: computed tomography
ELISA: enzyme-linked immunosorbent assay
FOV: field of view
HD: high definition
HSDI: high-speed digital imaging
PCI-e: Peripheral Component Interconnect Express
RGB: red-green-blue
SLM: selective laser melting

Edited by L Buis; submitted 17.11.20; peer-reviewed by W Choi, SL Lee; comments to author 09.02.21; revised version received 17.02.21; accepted 13.05.21; published 18.06.21.

Please cite as:
Kim Y, Oh J, Choi SH, Jung A, Lee JG, Lee YS, Kim JK
A Portable Smartphone-Based Laryngoscope System for High-Speed Vocal Cord Imaging of Patients With Throat Disorders: Instrument Validation Study
JMIR Mhealth Uhealth 2021;9(6):e25816
URL: https://mhealth.jmir.org/2021/6/e25816
doi:10.2196/25816
PMID:34142978

Original Paper

Web-Based Self-management Program (SPACE for COPD) for Individuals Hospitalized With an Acute Exacerbation of Chronic Obstructive Pulmonary Disease: Nonrandomized Feasibility Trial of Acceptability

Linzy Houchen-Wolloff[1,2], PhD; Mark Orme[1], PhD; Amy Barradell[1], MSc; Lisa Clinch[1], BSc; Emma Chaplin[1], BSc; Nikki Gardiner[1], MSc; Sally J Singh[1], PhD

[1]Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom
[2]Department of Respiratory Sceinces, University of Leicester, Leicester, United Kingdom

Corresponding Author:
Linzy Houchen-Wolloff, PhD
Centre for Exercise and Rehabilitation Science
University Hospitals of Leicester NHS Trust
Glenfield Hospital
Leicester
United Kingdom
Phone: 44 116 250 2759
Email: Linzy.Houchen@uhl-tr.nhs.uk

Abstract

Background: Hospital admissions due to the acute exacerbation of chronic obstructive pulmonary disease (AECOPD) are costly for individuals and health services. Pulmonary rehabilitation (PR) is known to reduce hospital readmissions when delivered after hospitalization, but the uptake and completion of PR following hospitalization remains poor (<10% of those eligible in the UK audit data). A web-based platform of the SPACE (Self-management Program of Activity Coping and Education) for COPD (chronic obstructive pulmonary disease) has previously shown promising results in patients with stable COPD but has not been tested following an AECOPD.

Objective: This study aims to assess the feasibility and acceptability of a web-based self-management program.

Methods: A nonrandomized feasibility study for patients with confirmed AECOPD who were deemed web literate was conducted. All patients consented during their hospitalization and received access to the website following discharge in addition to usual care. The program aims to facilitate patients to better understand and manage their condition through education and home-based exercises. Participants were asked to complete the Bristol COPD Knowledge Questionnaire at baseline and after 6 months. A total of 14 participants were also interviewed (n=8 completers; n=6 noncompleters) regarding their experiences with the web-based program and trial. The interviews were analyzed using thematic analysis.

Results: In total, 2080 patients were screened for eligibility, of which 100 patients (age: mean 71.2 years, SD 9.3 years; male: 55/100, 55%; forced expiratory volume in 1 second/forced vital capacity ratio: mean 0.46, SD 0.14; pack-years: mean 50.2, SD 31.0; current smokers: 35/100, 35%) were recruited (4.8% of those screened). The main reason for ineligibility was a lack of web literacy (1366/1980, 68.98%). In total, 18% (18/100) of patients had completed the web program by 6 months, with others still registered in the program (27/100, 27%), and more than half did not register (55/100, 55%). There was a mean change in Bristol COPD Knowledge Questionnaire scores at 6 months of 7.8 (SD 10.2) points. Qualitative interviews identified three main themes: preparing for, engagement with, and benefits of the study and program. A total of 57% (57/100) accepted a referral to PR on discharge and 19% (19/100) had completed the program after 6 months.

Conclusions: On the basis of the challenges of recruiting, retaining, and engaging participants in a web-based self-management program, it is not a feasible approach to roll out widely. This study acknowledges that this is a challenging time for patients with an AECOPD to engage in exercise and self-management education. However, for patients who were able to engage in such an intervention, the completion rate of PR was double the previous audit estimates from the United Kingdom, disease knowledge improved, and the intervention was of value to patients.

Trial Registration: ISRCTN Registry 13081008; https://www.isrctn.com/ISRCTN13081008

(JMIR Mhealth Uhealth 2021;9(6):e21728) doi:10.2196/21728

KEYWORDS

COPD; telehealth; digital health; internet; rehabilitation; quantitative; qualitative; exercise

Introduction

Background

Hospital admissions for the acute exacerbation of chronic obstructive pulmonary disease (AECOPD) represent a huge burden to the individual in terms of muscle dysfunction, breathlessness, and inactivity [1]. Furthermore, the AECOPD is costly to health services, particularly when an inpatient stay is required [2,3]. Pulmonary rehabilitation (PR) is a high-value and cost-effective intervention that is safe and improves exercise capacity and quality of life [4], outplaying pharmacotherapy and telehealth [5], and may offer a survival advantage for those who complete the program [6]. As such, PR is recommended by national and international guidelines soon after an exacerbation [7,8].

Despite this guidance and the established benefits of PR, access to, uptake to, and completion of postexacerbation rehabilitation are very poor. In the United Kingdom, less than 10% of all hospital discharges for the AECOPD completed posthospitalization PR [9]. More recently, the UK National Audit found that only 3% of the audit caseload is for postexacerbation rehabilitation, with the rest attributed to patients with stable respiratory disease. We also know from data from the United Kingdom that often the suitability for rehabilitation is not assessed at discharge by the clinician in as many as 44% of cases [10]. Therefore, many potentially suitable patients are missing from this valuable intervention. This problem is not unique to the United Kingdom; indeed, recent figures from the United States suggest that only 1.9% of patients hospitalized for chronic obstructive pulmonary disease (COPD) exacerbation receive PR within 6 months of discharge. The rate of uptake varies widely according to geographic region and ethnicity [11].

The reasons for nonuptake and poor completion are underpinned by complex reasons; some are related to the organization and system of delivery and others to patients' individual choices [12,13]. A key problem in the postexacerbation phase is that patients feel too unwell or breathless to attend a hospital or community program [14]. At the time of starting this study, there had been no randomized controlled trials (RCTs) of interventions to increase the uptake of early rehabilitation following exacerbation. One quasi-randomized study (n=115) with a high risk of bias indicated greater program completion and attendance rates in participants allocated to PR alongside a tablet computer (support for exercise training) compared with controls (PR only [15]). Other studies have explored the role of changing the timing of postexacerbation rehabilitation by moving this into the periexacerbation phase (inpatient [16]) or by delaying the start for 7 weeks [17] or 6 months [18]. These initiatives have had little benefit over control conditions, and recruitment was a challenge for delayed studies [17,18]. One feasibility study has also looked to reduce sedentary time in those hospitalized with an AECOPD using wearable technology (a vibration prompt to move at intervals throughout the day) for 2 weeks postdischarge [19]. Collectively, patients responded to 32.6% (106/325) of vibration prompts from the waist-worn device. Qualitative interviews indicated that being unwell and overwhelmed after an exacerbation was the main reason for not engaging with the intervention, and retention in this study was poor (52%).

To address the problem of uptake, over the past 10 years, we have developed home-based alternatives to attending a traditional center or community-based PR. These remote models have become increasingly relevant in the era of COVID-19, as the social distancing measures taken in many countries to suppress transmission of SARS-CoV-2 have had an immediate and profound effect on the provision of PR services [20]. SPACE (Self-management Program of Activity Coping and Education) for COPD is a self-management program of activity, coping, and education, which was coproduced as a 4-stage manual by health care professionals and patients [21]. In a series of studies, it has been shown to improve symptoms and exercise tolerance above usual care control groups and is noninferior to PR for improvements in quality of life [22,23]. When delivered in a hospital, SPACE for COPD was able to improve the quality of life and readiness for home above the control [24]. More recently, we have transitioned the SPACE for COPD program to a web-based format because there is a real ambition for health care services to engage with new technologies [25]. This has been driven by data showing that internet use in the >75-year age group is rising rapidly, closing the gap in the younger age groups [26]. Clearly, these would be the target age groups for rehabilitation interventions. To this end, we tested the web-based version of SPACE for COPD in secondary care and found the approach to be feasible and acceptable when compared with standard rehabilitation [27]. However, we have not yet tested the web-based version of SPACE for COPD in an acute, hospitalized population. We also know that patients are very inactive upon hospital discharge [28]; therefore, home-based solutions that act as a stepping stone to outpatient PR may be warranted. Otherwise, there is a drastic increase in patient expectations.

Objective

The primary aim of this study is to assess the feasibility and acceptability of a web-based program for individuals hospitalized with COPD exacerbation.

Methods

Population

This was a single-center, nonrandomized feasibility study. All patients admitted with an exacerbation of COPD to Glenfield Hospital, Leicester, were screened for eligibility by the specialist COPD nursing team and given at least 24 hours to consider the information. We included patients with an email address who were web literate (used a tablet, PC, laptop, or device at least

once per week). To assess this, we asked patients about the types of devices used, the time spent on the web, and the types of web-based activities for each patient (eg, emailing, internet banking, and web-based shopping). A decision on web literacy was at the discretion of the recruiting clinician. The program was predominately delivered via a tablet. Patients could use their own device or borrow an Android tablet, for instance, if they usually borrowed a family member or did not have access for the duration of the study. Borrowed devices were locked (SureLock software, 42Gears Mobility Systems Limited), other than for access to the SPACE for COPD website. People with significant neuromuscular or cardiovascular comorbidities limiting physical activity (typical exclusion criteria for PR) and those who were unable to read and write in English were excluded. Currently, the website is available only in the English language.

Intervention

Patients were given access to the SPACE for COPD program as an inpatient. A passport card with log-in information and staff contact details and a user manual were given to the patient on discharge along with a verbal introduction to the program (alongside viewing this on a tablet) by the COPD nursing team.

SPACE for COPD is an interactive web-based program that offers a comprehensive package of exercise and self-management education. The program was structured to guide the user through four stages, each of which has specific tasks that the user needs to achieve before progressing to the next stage. Tasks included creating and updating their own short-term goals, completing knowledge tests on COPD and exercising safely, and reading or watching videos on specific topics, such as inhaler techniques or healthy eating. The program was described in detail by Chaplin et al [27]. In stage 2, patients were asked to record their aerobic walking exercise, and for this study, we devised a symptom diary that linked to the patients' individual exacerbation action plan. The web-based program usually takes approximately 11 weeks to complete for patients with stable COPD [27], although this patient cohort had access for 1 year, to promote long-term behavior change and maintenance. However, the outcomes were assessed at 6 months. Special features of the program include videoconferencing (where patients could have a live consultation with the COPD nursing team at an allocated time), a moderated blog section (where patients could share their experiences with others), and an *ask the expert* facility (to email the COPD nursing team). The *ask the expert* emails were monitored by the specialist COPD nursing team during working hours (Monday to Friday, 8 AM-4 PM). Prompts to log on to the website and record activity were automatically generated by the program and sent via email if patients failed to record activity for 7 consecutive days. Patients also received a telephone call from specialist COPD nurses within the first 5 days following discharge (as is usual care).

Usual care, including referral to and attendance at PR, was not affected by this trial; patients received a telephone call from the COPD nursing team within 5 days of discharge and had a scheduled follow-up appointment with a consultant within 3 months of discharge.

Outcomes

The primary outcome was the feasibility of the intervention (uptake to the intervention: percentage of patients recruited out of the total number screened and completion rates of the Bristol COPD Knowledge Questionnaire [BCKQ]). This questionnaire was chosen because the authors felt that it was the least likely to be influenced by illness and natural recovery following the AECOPD. The secondary outcomes were the acceptability of the intervention and trial (qualitative interviews), intervention engagement (web usage statistics: number of log-ins and use of web features captured directly from the administrator section of the website), and uptake to outpatient PR (uptake and completion rates in those referred). All outcomes were assessed at baseline (in hospital upon enrollment to the study) and 6 months following enrollment in the study, regardless of whether the patient had engaged with the web-based program or PR during the 6-month period.

Analysis

Quantitative Analysis

Data are described as mean (SD), median (IQR), or frequency (%), as appropriate. No inferential statistics were obtained owing to the feasibility nature of the trial. As is convention with feasibility studies, a formal sample size is not required [29]. A total of 100 patients were recruited.

Qualitative Analysis

To measure patients' views on the acceptability of the web-based program and the study, qualitative interviews were conducted with completers and dropouts on a purposive sampling basis (completer and noncompleter interview schedules are shown in Multimedia Appendices 1 and 2 respectively). Patients who agreed to be contacted for an interview were approached following the 6-month study time point, as it was anticipated that most patients would have accessed and/or completed the web-based program by this point.

The interviews were conducted by the qualitative researcher AB, who was relatively naïve to the web-based program and study processes. After each interview, the researcher wrote reflective and methodological notes to assist the analytic process and enhance the rigor of the results.

The interviews were audio recorded and transcribed verbatim using an external source. Analysis of the interviews was conducted manually using the thematic analysis framework by Braun and Clarke [30]. The six phases of the framework were followed by AB, LHW, SJS, and MO, who independently coded the interviews. These phases included data familiarization, data coding, searching for themes, reviewing themes, defining themes, and composing the narrative. Agreement of themes was made by the four coders.

Results

Primary Outcome

A total of 2080 patients were screened over 2 years (May 2015 to September 2017) to obtain a sample of 100 patients. The proportion of patients recruited as 4.8% of those screened

(100/2080). The predominant reason for exclusion in approximately (1366/1980, 68.98%) of cases was that patients were not web literate or did not have an email address. Table 1 provides the reasons for exclusion or nonuptake.

Table 1. Reasons for exclusion or nonuptake (N=1980).

Reason	Patients, n (%)
Not web literate, no email address	1366 (68.98)
Unwilling	297 (15)
Comorbidities precluding involvement in the study	238 (12.02)
Has done pulmonary rehabilitation or SPACE[a] previously (did not want to do it again)	40 (2.02)
On another research study	20 (1.01)
Unable to read English	19 (0.95)

[a]SPACE: Self-management Program of Activity Coping and Education.

Of the participants recruited, the mean age of participants was 71 (SD 9) years; they had severe disease and a mean smoking pack-year history of 50.2 (SD 31.0) years (Table 2). There was a good split between male and female participants: 55% (55/100) males, 35% (35/100) were current smokers, and most had at least one other comorbidity (93/100, 93%).

Table 2. Baseline characteristics of recruited participants.

Variables	Value, mean (SD)
Age (years)	71.2 (9.3)
BMI	28.1 (9.8)
FEV_1[a]/FVC[b]	46.2 (13.9)
FEV_1 (% predicted)	44.8 (18.3)
Pack-years	50.2 (31.0)

[a]FEV_1: forced expiratory volume in 1 second.

[b]FVC: forced vital capacity.

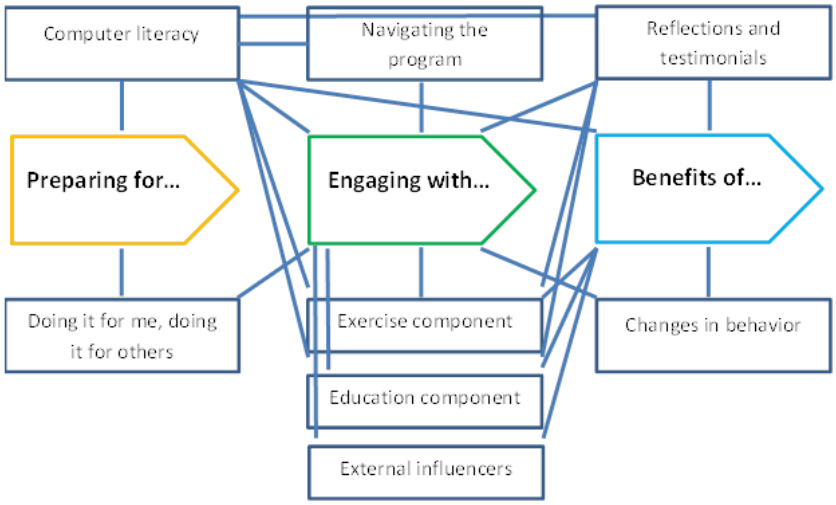
Secondary Outcomes

Disease Knowledge

The change in the BCKQ score was 7.8 (SD 10.2) points, an increase of 21% (prescreening score: mean 37.1, SD 9.5; postscreening score: mean 44.9, SD 9.4). This was done in 42 patients who returned the questionnaires at 6 months.

Qualitative interviews with a sample of 14 patients (n=8 completers; n=6 noncompleters) identified three themes, with a total of eight subthemes. The relationships between the themes are presented in Figure 1. The identified themes were *preparing for*, *engaging with*, and *benefits of* the web-based program.

Figure 1. Qualitative thematic map.



Preparing for: Doing It for Me

The patients described a number of reasons why the web-based program incited their internal motivation to engage with a telerehabilitation program. Some studies have described the fear they experienced from a recent COPD exacerbation as a catalyst:

> *...when you're recovering from something that frightens you to death, you say yes to everything. So that's why I said yes.* [Participant 69, completer]

They also felt that the program offered an opportunity to learn how to manage their condition more successfully through the ongoing support of health care professionals:

> *I wanted to learn more and to see if there was anything that would be beneficial to the breathing obviously. Because I mean we know that the rehabilitation works. I can vouch for that. Unfortunately it has a very short shelf life unless you continue it. So I was hoping that just get information because there really isn't that much information. Your own GP doesn't seem to have, I mean I've got a fantastic GP, but he hasn't got a lot of information on COPD. If you have a problem, he'll refer you to the respiratory nurse who comes about once a year.* [Participant 39, completer]

The appeal of a home-based program was also evident among patients. They felt it offered them the ability to complete the program at their own pace, and this flexibility allowed them to fit it into their established routines:

> *...she said you can do it at home on your laptop and in my own time. I haven't got to go out the house, well I've got to go out the house to walk.* [Participant 100, noncompleter]

Preparing for: Doing It for Others

A subset of patients described external influences as factors that motivated them to engage in the program. Some felt that the reassurance health care professionals offered was what they needed to engage with the program. Others felt that their families were instrumental in encouraging their engagement with the program. One patient described how the support of her family gave her the confidence to engage:

> *but I thought—like my children said to me, if you're not doing that mam, what are you going to be doing. Thought well, yeah, they're right, why shouldn't I do it.* [Participant 52, completer]

For another, the program was viewed as an opportunity to share the learning materials with their loved ones in an effort to improve both their health:

> *...I was hoping that at the end of it I would have a bit more information as to - because my missus had got COPD as well and there was something we could learn from it.* [Participant 39, completer]

Preparing for: Computer Literacy Skills and Suitability

Overall, patients felt that their computer literacy skills were vital to their ability to engage with the program. Despite the assessment of computer literacy skills before enrollment in the study, some patients felt that their age or generation acted as a barrier to their ability and many lacked confidence in their skills:

> *I mean for you and for your generation, you don't even think about it do you, you just do it. For us, we never had it. It was chalk and slate!* [Participant 39, completer]

Engaging With: Navigating the Program

All patients felt that the program was complex in nature and learning to navigate it felt like a steep learning curve. Some felt it was too steep a learning curve for their computer literacy skills, which resulted in reduced motivation to engage with the program or, for some, to disengage from the program entirely:

> *...you spend more time on the computer than walking. So that defeated the object with me anyway.* [Participant 100, noncompleter]

To tackle this issue, some patients sought help from family members or health care professionals but still found the interface challenging and at times overwhelming:

> *It wouldn't let me in. I tried it on Safari and it kept refusing me and then I finally got on...And I just couldn't get on with it. I got it on in the end with the help from my daughter and then all the questionnaires, never ending, just one page after another...what's your lung capacity? How do I know? What is it when you breathe in, what is it when you breathe out?—terrible!* [Participant 100, noncompleter]

For those patients who persevered with the program, they felt self-discipline, self-motivation, and allocating time to learn how to navigate the program were essential to their progression:

> *As I say, after about three to four weeks I was very confident or confident, I've never been very confident at any of it, confident, but I must be honest it did take me about that long before I felt happy doing it. You know, I did it and I'd sit there for ages doing it and my missus would say to me have you done it yet? But, you know, you'd persevere....* [Participant 39, completer]

Some patients expressed frustration due to the interaction they had with the program, particularly the email reminders to exercise daily as they felt it lacked the acknowledgment of their individual internal and external barriers to exercise:

> *I found myself getting a bit wound up when I'd not carried out certain tasks to the satisfaction of the, you know, the interaction was you haven't exercised this week. No I haven't because I've not been well!* [Participant 69, completer]

Other patients embraced the exercise reminders and enjoyed the *little nudge*, as it helped them develop exercise habits:

> *I found it very good because you're doing it at your own pace. And it wasn't excessive and it just gave you that little nudge in the back of your mind. You'd look at your emails and it would come up SPACE and*

> *you'd think oh ah. It were more like, not a guilt trip, that's wrong, but it just give you that little nudge and you think. But once you got into it, you know, you did it automatic which was a good thing.* [Participant 39, completer]

A further facility to initiative interaction with health care professionals was the teleconferencing option, but this was not used by any participant. Some patients felt it was too advanced for them, others felt it lacked a personal approach from staff, and some were unaware of this option:

> *I think I did see it, but that, I'm not, I'm better off person to person, not person to screen.* [Participant 78, noncompleter]

Engaging With: Exercise Component

For patients who reached the exercise component of the program (stage 2), many felt goal setting was beneficial, as it motivated them to exercise more, which in turn had a positive impact on their quality of life:

> *Yeah, my breathing is better when I've done the exercises, when I've done the walking. It don't feel at the time, I'm gasping like hell. Later in the day you feel the benefits of it. It always clears your lungs out, it clears your lungs doing the walk. And I enjoy that.* [Participant 100, noncompleter]

The daily recording and monitoring of the patients completed also helped some to identify patterns in their exercise behavior, and from these patterns, they could identify earlier onset of COPD exacerbations:

> *And what was interesting with that was if you were honest and put it down on the programme, you started to see a pattern building up. And I don't know what the pattern meant or whatever but you could see when you were coming up to, you know, one of the and you needed the antibiotics, you could see it before.* [Participant 39, completer]

However, patients felt that their ability to exercise was restricted by factors outside their control. For example, some did not prioritize exercise if they felt unwell or if the weather was deemed inclement. However, one patient who faced this scenario did not let this deter him from exercising:

> *I go on my bike when I can't go out then I'll stand and do a few knee exercises, stretches and bending and things.* [Participant 100, noncompleter]

Engaging With: Education Component

Overall, patients felt that the education component developed their understanding of COPD and increased their knowledge and skills regarding self-management. They felt that the program highlighted the importance of exercise and gave them important practical skills:

> *The breathing, the way that you breathe. The necessity to exercise is the thing that comes out of all of it, even if it's only small amounts. And if you can't hit what you set yourself to do, don't beat yourself up with it. Just do it, do the best you can, and that's the thing that came out of it. If you can't do what you want to do on that particular day, do what you can do.* [Participant 39, completer]

Some patients felt that the program also helped them understand the relationship between COPD and mood. This understanding helped them recognize their feelings as a natural reaction to living with a long-term condition:

> *...one of the other things that came out about it was things like getting depressed. I don't tend to get depressed, I tend to get fed up with it all and you think is this it, whatever? Which is again very normal but you don't know that and you think what's happening? They tell you that.* [Participant 39, completer]

Engaging With: External Influencers

Some patients felt that their ability to complete the program was affected by factors outside of their control, which reduced their motivation and level of discipline. Exacerbations of COPD, comorbidities, and becoming a caregiver for a loved one were all described as reasons for noncompletion:

> *So once I got used to it I didn't really have a problem with it. And I followed that religiously. And then we hit a snag, but it was nothing to do with the course. My missus had to go into hospital for three months because she had a brain haemorrhage. So when she came out, she still needed, we needed to put the time in to get her back to normal again, which was three, four, five months. So that then took a little bit of a backseat and I found then I couldn't quite catch up.* [Participant 39, completer]

Benefits of: Reflections and Testimonials

Patients who engaged in the program felt it gave them a focus and drive to take control of their health. Patients felt that the program helped them reassess their illness perceptions, which, in turn, made them feel less self-conscious, more confident in self-managing their COPD, and gave them hope for the future:

> *I feel in control of it. Instead of it controlling me* [Participant 52, completer]

Benefits of: Changes in Behavior

Some patients felt that the program inspired them to challenge their health behaviors and encouraged them to make positive changes. Patients attributed increases in daily exercise to the program, and others found affirmation for smoking cessation:

> *[I've learned] Never to smoke again.* [Participant 88, completer]

Program completers and noncompleters all felt that the program helped them find enjoyment in exercise, and this encouraged a continuation of exercise following program completion. To facilitate this, some patients purchased exercise equipment, downloaded fitness apps, recorded their daily exercise, and began to embed exercise into their daily routine. This increase in exercise meant that some participants were now more capable of carrying out their daily activities:

> *I managed to do all my garden last year instead of having a gardener...I can now have a shower and I wash everywhere down.* [Participant 52, completer]

User Statistics

None of the participants used videoconferencing or blog facilities. The median number of log-ins per person was 14 and emails to the team was 1. Patients generally recorded more than one goal throughout the course of the program, an average of 3.6 (SD 2.4) per person, and an average of 16.5 (SD 32.7) of the exercise sessions were recorded.

At 6 months, 18% (18/100) of participants had completed the web-based program, 27% (27/100) were still registered, and 55% (55/100) had not registered, despite prompts by email and a phone call from the nursing team within the first 5 days following discharge. The fact that one-fourth were still registered at 6 months suggests that the acute population may need longer to complete the program compared with a stable population.

Uptake to PR

Of the total 100 participants in the sample, 57 accepted a referral for rehabilitation. Of these, 47 were assessed, and 35 started a program; 19% (19/100) of the total population completed either a hospital or community outpatient rehabilitation program.

Discussion

Principal Findings

On the basis of the challenges of recruiting, retaining, and engaging participants in the SPACE for COPD web-based self-management program, it is not a feasible approach to roll out widely. In the face of organizational pressure to provide health information on the web and despite statistics suggesting otherwise, digital literacy in this group was lower than anticipated, as evidenced by the low number of patients deemed to be web literate.

Engagement in the program was poor. At 6 months, only 18% (18/100) had completed the web-based program, 27% (27/100) were still registered, and 55% (55/100) had not registered, despite prompts via email and phone calls from the nursing team. The fact that one-fourth of participants were still registered at 6 months perhaps indicates that an acute population takes longer to complete the program compared with a stable population (stable—complete within 11 weeks [27]). The lack of engagement is surprising, given that patients willingly signed up to the study; it may be that patients overestimated their information technology skills. In a recent qualitative study by Slevin et al [31], patients reported a willingness to take a more active role in self-management using digital health technologies. They perceived digital health technologies as potentially enhancing their self-management skills *by improving self-efficacy and engagement and by supporting health care professionals to practice preventative care provision*. However, we are aware that, although web usage among people aged >75 years is increasing [26], this may range from simply sending an email through to more complex tasks such as web-based banking or navigating websites. For instance, older patients, those with a lower socioeconomic status and those with more severe health needs are less likely to use technology or to handle eHealth-based tasks [32]. This is of concern given that the PR community has been challenged to rapidly provide innovative and alternative ways of delivering rehabilitation in the face of the COVID-19 epidemic, despite the evidence that the efficacy of digital self-management and rehabilitation programs in COPD is uncertain [33]. It may be that for this population, additional training on using the website will be required with a competency assessment built in to increase uptake, engagement, and completion.

The patient interviews revealed a strong divide among participants' willingness to engage in the program. Those who engaged demonstrated competent computer literacy skills, attributed more value to the program, experienced more external motivation from it, and were able to use it flexibly to adapt to their health and other life events. Those who did not engage with the program experienced greater difficulty in navigating the website, which they felt reduced their internal and external motivation to complete the program. This finding was also true of the TELEKAT (Telehomecare, Chronic Patients and the Integrated Healthcare System) study [34], in which severe to very severe patients with COPD expressed commitment to the program if they had a prior interest in telehealth or new technologies. However, this commitment did waiver when patients experienced COPD flare-ups. Therefore, it would be helpful to identify patients who would benefit from a different approach, particularly around the time of an exacerbation.

These results suggest that patients' exercise behavior was influenced by their intention to exercise, and this was further negated by their internal and external motivation, their opportunity to exercise, and their physical ability to exercise. These findings support the *capability, opportunity, motivation, and behavior* model of behavior change [35], which describes how behavior is driven by intention, which is further negated by motivation, opportunity, and capability. The model predicts that when one or more of these factors are reduced, the intention to carry out the behavior is also reduced.

However, in those recruited to this study, the completion rate of PR was 19%. This is double the previous audit estimates from the United Kingdom and significantly higher than the 1.9% proposed recently from the United States [9,11]. In addition, disease knowledge improved by 21% in this cohort. Although there is no accepted minimum clinically important difference for BCKQ, typical changes following outpatient PR are in the region of 18% [36].

Therefore, we would conclude that web-based strategies may be a viable stepping stone to postexacerbation PR in those able and willing to engage with the program.

Limitations

The main limitation of the study is that it was not an RCT; therefore, the effects of natural recovery were not considered. However, we chose to look at disease knowledge as a secondary outcome (rather than an outcome such as exercise tolerance or muscle strength), which is less likely to be influenced by exacerbation recovery. Furthermore, as this was a feasibility study, we did not intend to infer clinical effectiveness. We

accept that postal returns of questionnaires are not the most reliable way to return data, and we followed this up with a phone call prompt to increase data completion. It may have been interesting to examine the differences in outcomes such as hospital readmission and uptake to PR in a matched usual care–only cohort (ie, those not receiving the web program), but we did not seek ethical approval to do this.

This was also a single-center study, which may introduce selection bias and decrease the generalizability of the findings to other settings. Further concordance of qualitative themes between the four coders using statistical methods (eg, κ) would have strengthened these data further.

As stated previously, patients or recruiting clinicians may have overestimated the patients' ability to navigate the web program. This was despite strictly adhering to our screening criteria (must be web literate and must have a valid email address), a validated digital literacy measure would have been useful to aid screening.

The timing of enrollment in the study and introduction of the intervention were pragmatic and not standardized. We tried to do this as close to discharge as possible, but this was not always easy to predict in practice. We appreciate that a hospital admission may not be the best time for patients to be receptive to new information. In particular, it appears that cognitive function is specifically impaired during exacerbation but may recover [37]. Therefore, shifting our intervention introduction to a later time point following discharge may have been preferable.

Although the completion of postexacerbation rehabilitation in this study is impressive, it is of course with the caveat that these were self-selecting, enthusiastic research participants rather than *all comers*. Therefore, it is likely that these participants had a personal interest in self-management of their disease. In addition, because the web-based program acted as a graduated transition into outpatient PR, there may have been participants who stopped using the web-based program when they started PR or vice versa (ie, used the web program and did not attend PR). We have not been able to tease out these nuances in our data.

Comparison With Prior Work

Several gaps still exist in the literature on the topic of increasing access to and participation in PR. There may be instances where completion rates for PR are buried within manuscripts as a secondary outcome, but as it is not the primary outcome, it is unlikely to change guidance. This may be because we have no idea what an acceptable level of uptake or completion would be; therefore, studies tend to choose primary outcomes based on the quality of life or exercise measures, where we already have a wealth of data or minimum clinically important difference values. Therefore, high-quality research is needed to review complex interventions with uptake or completion as the primary outcome. In addition, we might assume that many centers have adapted their programs in pragmatic service improvement initiatives to increase acceptability without robust testing of these methods. This would account for the gaps in the literature.

Since we have completed our study, one RCT has been published, which evaluated the effect of a co-designed (by patients and health care professionals) education video as an adjunct to usual care on posthospitalization PR uptake. PR uptake was 41% and 34% in the usual care and intervention groups, respectively (P=.37), with no differences in secondary (PR referral and completion) or safety (readmissions and death) outcomes. Unfortunately, 40% (6/15) of participants interviewed did not recall receiving the video.

Other digital health apps or telerehabilitation have been reported in the literature for COPD populations [38,39]. Of note, the myCOPD app is a web-based 6-week rehabilitation program, which was found to be noninferior to a conventional PR delivered in face-to-face sessions in terms of the effects on walking distance and symptom scores (COPD assessment test) at 7 weeks [38]. The program was safe and well tolerated; however, it is worth noting that the population recruited was milder with fewer comorbidities than is usually reported in PR studies and that adherence to exercise sessions was slightly lower in the web-based group than in the face-to-face sessions per week [38]. The same app has been routinely offered to stable patients in Hammersmith and Fulham Respiratory Clinics (London, United Kingdom). Although two-thirds of patients were eligible to use the app (64%, 163/253), this has not translated into uptake (56%); 15% (297/1980) of our population declined to take part or engage with the PR program (10%), and this challenges the assumption that this digital app can be delivered as a suitable alternative to standard PR [40]. It is not clear what engagement referred to in this report, but if it was the completion of the app program, then a 10% noncompletion rate is disappointing and concurs with poor completion of the web program at 6 months in this study (18/100, 18%). Recently, Polgar et al [41] assessed the digital habits of patients referred to PR during the COVID-19 pandemic. There was significant heterogeneity in access to and confidence in using the internet, with 31% having never previously accessed the internet, 48% confident using the internet, and 29% reporting no interest in accessing any component of PR through a web-based application [41].

A recent qualitative work by Janaudis-Ferreira et al [42] has begun to *set the stage* to design a more acceptable PR program following an exacerbation of COPD. In this study, one-on-one interviews were conducted to explore the views of 13 patients and 11 health care professionals on PR after the AECOPD and how participation could be enhanced. Four main themes were identified: (1) uncertainty regarding the timing of PR; (2) tailored and flexible manner to deliver PR programs with a gradual start; (3) education for all; and (4) logistical, disease-related, and psychological barriers. Theme 2 is particularly interesting and aligns well with this study, as the web-based intervention may be thought of as a *bridge* to start a formal PR program. Theme 4 chimes with some of the participant quotes from our qualitative work.

Conclusions

On the basis of the challenges of recruiting, retaining, and engaging participants in this web-based self-management program, the SPACE for COPD web-based self-management program is not a feasible approach to roll out widely following an AECOPD. It appears that the COPD population may not be

equipped and ready for digital self-management interventions following an AECOPD, without additional training or support. This work acknowledges that this is a challenging time for patients with an AECOPD to engage in exercise and self-management education. However, for patients able to engage with such an intervention, the completion of PR was double the previous estimates from UK audit or research data, disease knowledge improved, and it encouraged positive behavior change and was of value to patients. Therefore, with further refinement, web-based (or other home-based) strategies may be a viable stepping stone for PR. Identifying the patients most likely to benefit from such strategies is warranted, particularly if social distancing measures for COVID-19 are to continue.

Acknowledgments

The authors would like to acknowledge the help of the specialist COPD nursing team at Glenfield Hospital, Leicester, who identified suitable participants for the trial and monitored the patients' progress on the web-based program and Sally Schreder, Pratiksha Kanabar, Agnieszka Glab, Chris Brough, and Stacey Hewitt for trial setup, data collection, and data entry assistance. Finally, the authors thank the PR team at Glenfield Hospital, Leicester, who provided the PR uptake and completion data. This work was funded by the East Midlands Academic Health Science Network Funding Phase 2. This work was supported by the National Institute of Health Research Leicester Biomedical Research Center. The views expressed are those of the authors and not necessarily those of the National Health Service, National Institute of Health Research, or Department of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Completers' interview schedule.
[DOC File , 264 KB - mhealth_v9i6e21728_app1.doc]

Multimedia Appendix 2
Noncompleters' interview schedule.
[DOC File , 264 KB - mhealth_v9i6e21728_app2.doc]

References

1. Jones SE, Barker RE, Nolan CM, Patel S, Maddocks M, Man WD. Pulmonary rehabilitation in patients with an acute exacerbation of chronic obstructive pulmonary disease. J Thorac Dis 2018 May;10(Suppl 12):1390-1399 [FREE Full text] [doi: 10.21037/jtd.2018.03.18] [Medline: 29928521]
2. Britton M. The burden of COPD in the U.K.: results from the confronting COPD survey. Respir Med 2003 Mar;97:71-79. [doi: 10.1016/s0954-6111(03)80027-6]
3. Perera PN, Armstrong EP, Sherrill DL, Skrepnek GH. Acute exacerbations of COPD in the United States: inpatient burden and predictors of costs and mortality. COPD 2012 Apr 12;9(2):131-141. [doi: 10.3109/15412555.2011.650239] [Medline: 22409371]
4. Puhan MA, Gimeno-Santos E, Cates C, Troosters T. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2016 Dec 08;12:CD005305 [FREE Full text] [doi: 10.1002/14651858.CD005305.pub4] [Medline: 27930803]
5. Mak V. Treatment guidelines for COPD - going for GOLD? Primary Care Respiratory UPDATE. 2017. URL: https://www.pcrs-uk.org/sites/pcrs-uk.org/files/GoingforGOLD_PCRUSept2017.pdf [accessed 2021-04-24]
6. Houchen-Wolloff L, Williams J, Green R, Woltmann G, Steiner M, Sewell L, et al. Survival following pulmonary rehabilitation in patients with COPD: the effect of program completion and change in incremental shuttle walking test distance. COPD 2017 Dec;Volume 13:37-44. [doi: 10.2147/copd.s143101]
7. Bolton CE, Bevan-Smith EF, Blakey JD, Crowe P, Elkin SL, Garrod R, British Thoracic Society Pulmonary Rehabilitation Guideline Development Group, British Thoracic Society Standards of Care Committee. British Thoracic Society guideline on pulmonary rehabilitation in adults. Thorax 2013 Sep 23;68 Suppl 2(Suppl 2):ii1-i30. [doi: 10.1136/thoraxjnl-2013-203808] [Medline: 23880483]
8. Chronic obstructive pulmonary disease in over 16s: diagnosis and management. NICE guideline [NG115]. 2018. URL: https://www.nice.org.uk/guidance/ng115 [accessed 2021-04-24]
9. Jones SE, Green SA, Clark AL, Dickson MJ, Nolan A, Moloney C, et al. Pulmonary rehabilitation following hospitalisation for acute exacerbation of COPD: referrals, uptake and adherence. Thorax 2014 Feb 14;69(2):181-182. [doi: 10.1136/thoraxjnl-2013-204227] [Medline: 23945168]

10. Pulmonary rehabilitation: an exercise in improvement – combined clinical and organisational audit 2017. Royal College of Physicians. 2018. URL: https://www.rcplondon.ac.uk/projects/outputs/pulmonary-rehabilitation-exercise-improvement-combined-clinical-and-organisational [accessed 2021-04-24]
11. Spitzer KA, Stefan MS, Priya A, Pack QR, Pekow PS, Lagu T, et al. Participation in pulmonary rehabilitation after hospitalization for chronic obstructive pulmonary disease among medicare beneficiaries. Annals ATS 2019 Jan;16(1):99-106. [doi: 10.1513/annalsats.201805-332oc]
12. Benzo R, Wetzstein M, Neuenfeldt P, McEvoy C. Implementation of physical activity programs after COPD hospitalizations: lessons from a randomized study. Chron Respir Dis 2015 Feb 15;12(1):5-10 [FREE Full text] [doi: 10.1177/1479972314562208] [Medline: 25511306]
13. Milner K, O'Connor M. Shared decision making and decision aids: an important part of evidence-based practice. J Nurs Educ 2017 Nov 01;56(11):702-703. [doi: 10.3928/01484834-20171020-14] [Medline: 29091244]
14. Harrison SL, Robertson N, Apps L, Steiner MC, Morgan MD, Singh SJ. "We are not worthy"--understanding why patients decline pulmonary rehabilitation following an acute exacerbation of COPD. Disabil Rehabil 2015 Jul 10;37(9):750-756. [doi: 10.3109/09638288.2014.939770] [Medline: 25009949]
15. Ringbaek TJ, Lavesen M, Lange P. Tablet computers to support outpatient pulmonary rehabilitation in patients with COPD. Eur Clin Respir J 2016 May 24;3(1):31016 [FREE Full text] [doi: 10.3402/ecrj.v3.31016] [Medline: 27225598]
16. Greening NJ, Williams JE, Hussain SF, Harvey-Dunstan TC, Bankart MJ, Chaplin EJ, et al. An early rehabilitation intervention to enhance recovery during hospital admission for an exacerbation of chronic respiratory disease: randomised controlled trial. Br Med J 2014 Jul 08;349:g4315 [FREE Full text] [doi: 10.1136/bmj.g4315] [Medline: 25004917]
17. Revitt O, Sewell L, Singh S. Early versus delayed pulmonary rehabilitation: a randomized controlled trial - can we do it? Chron Respir Dis 2018 Aug 22;15(3):323-326 [FREE Full text] [doi: 10.1177/1479972318757469] [Medline: 29471678]
18. Puhan MA, Spaar A, Frey M, Turk A, Brändli O, Ritscher D, et al. Early versus late pulmonary rehabilitation in chronic obstructive pulmonary disease patients with acute exacerbations: a randomized trial. Respiration 2012 Aug 16;83(6):499-506 [FREE Full text] [doi: 10.1159/000329884] [Medline: 21846962]
19. Orme MW, Weedon AE, Saukko PM, Esliger DW, Morgan MD, Steiner MC, et al. Findings of the Chronic Obstructive Pulmonary Disease-Sitting and Exacerbations Trial (COPD-SEAT) in reducing sedentary time using wearable and mobile technologies with educational support: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e84 [FREE Full text] [doi: 10.2196/mhealth.9398] [Medline: 29643055]
20. Houchen-Wolloff L, Steiner MC. Pulmonary rehabilitation at a time of social distancing: prime time for tele-rehabilitation? Thorax 2020 Jun 12;75(6):446-447. [doi: 10.1136/thoraxjnl-2020-214788] [Medline: 32398319]
21. Apps LD, Mitchell KE, Harrison SL, Sewell L, Williams JE, Young HM, et al. The development and pilot testing of the self-management programme of activity, coping and education for chronic obstructive pulmonary disease (SPACE for COPD). Int J Chron Obstruct Pulmon Dis 2013 Jul;8:317-327 [FREE Full text] [doi: 10.2147/COPD.S40414] [Medline: 23874093]
22. Horton EJ, Mitchell KE, Johnson-Warrington V, Apps LD, Sewell L, Morgan M, et al. Comparison of a structured home-based rehabilitation programme with conventional supervised pulmonary rehabilitation: a randomised non-inferiority trial. Thorax 2018 Jan 29;73(1):29-36. [doi: 10.1136/thoraxjnl-2016-208506] [Medline: 28756402]
23. Mitchell KE, Johnson-Warrington V, Apps LD, Bankart J, Sewell L, Williams JE, et al. A self-management programme for COPD: a randomised controlled trial. Eur Respir J 2014 Dec 03;44(6):1538-1547 [FREE Full text] [doi: 10.1183/09031936.00047814] [Medline: 25186259]
24. Singh S, Rees K, Gelder C, Morgan M, Johnson-Warrington V. Can a supported self-management program for COPD upon hospital discharge reduce readmissions? A randomized controlled trial. COPD 2016 Jun:1161. [doi: 10.2147/copd.s91253]
25. The Topol Review. Preparing the healthcare workforce to deliver the digital future. The NHS Constitution. 2019. URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf [accessed 2021-04-24]
26. Internet use in the UK annual estimates by age, sex, disability and geographical location. Office for National Statistics. 2019. URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2019 [accessed 2021-04-24]
27. Chaplin E, Hewitt S, Apps L, Bankart J, Pulikottil-Jacob R, Boyce S, et al. Interactive web-based pulmonary rehabilitation programme: a randomised controlled feasibility trial. BMJ Open 2017 Mar 31;7(3):e013682 [FREE Full text] [doi: 10.1136/bmjopen-2016-013682] [Medline: 28363923]
28. Orme MW, Harvey-Dunstan TC, Boral I, Chaplin EJ, Hussain SF, Morgan MD, et al. Changes in physical activity during hospital admission for chronic respiratory disease. Respirology 2019 Jul 07;24(7):652-657 [FREE Full text] [doi: 10.1111/resp.13513] [Medline: 30845363]
29. Lancaster GA, Dodd S, Williamson P. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

31. Slevin P, Kessie T, Cullen J, Butler MW, Donnelly SC, Caulfield B. A qualitative study of chronic obstructive pulmonary disease patient perceptions of the barriers and facilitators to adopting digital health technology. Digit Health 2019 Aug 25;5:2055207619871729 [FREE Full text] [doi: 10.1177/2055207619871729] [Medline: 31489206]
32. Witry M, Comellas A, Simmering J, Polgreen P. The association between technology use and health status in a chronic obstructive pulmonary disease cohort: multi-method study. J Med Internet Res 2018 Apr 02;20(4):e125 [FREE Full text] [doi: 10.2196/jmir.9382] [Medline: 29610113]
33. Shaw G, Whelan ME, Armitage LC, Roberts N, Farmer AJ. Are COPD self-management mobile applications effective? A systematic review and meta-analysis. NPJ Prim Care Respir Med 2020 Apr 01;30(1):11 [FREE Full text] [doi: 10.1038/s41533-020-0167-1] [Medline: 32238810]
34. Dinesen B, Huniche L, Toft E. Attitudes of COPD patients towards tele-rehabilitation: a cross-sector case study. Int J Environ Res Public Health 2013 Nov 15;10(11):6184-6198 [FREE Full text] [doi: 10.3390/ijerph10116184] [Medline: 24247995]
35. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
36. White R, Walker P, Roberts S, Kalisky S, White P. Bristol COPD Knowledge Questionnaire (BCKQ): testing what we teach patients about COPD. Chron Respir Dis 2006 Sep 02;3(3):123-131 [FREE Full text] [doi: 10.1191/1479972306cd117oa] [Medline: 16916006]
37. Dodd JW, Getov SV, Jones PW. Cognitive function in COPD. Eur Respir J 2010 Apr 31;35(4):913-922 [FREE Full text] [doi: 10.1183/09031936.00125109] [Medline: 20356988]
38. Bourne S, DeVos R, North M, Chauhan A, Green B, Brown T, et al. Online versus face-to-face pulmonary rehabilitation for patients with chronic obstructive pulmonary disease: randomised controlled trial. BMJ Open 2017 Jul 17;7(7):e014580 [FREE Full text] [doi: 10.1136/bmjopen-2016-014580] [Medline: 28716786]
39. Bhatt SP, Patel SB, Anderson EM, Baugh D, Givens T, Schumann C, et al. Video telehealth pulmonary rehabilitation intervention in chronic obstructive pulmonary disease reduces 30-day readmissions. Am J Respir Crit Care Med 2019 Aug 15;200(4):511-513 [FREE Full text] [doi: 10.1164/rccm.201902-0314LE] [Medline: 30978302]
40. Lound A, O'Toole D, Spurway R, Killeen L. The accessibility and use of a digital health platform in a cohort of stable community COPD patients. Primary Care Respiratory Society (PCRS). 2019. URL: https://www.researchgate.net/publication/343191024_The_accessibility_and_use_of_a_digital_health_platform_in_a_cohort_of_stable_community_COPD_patients [accessed 2021-04-24]
41. Polgar O, Aljishi M, Barker RE, Patel S, Walsh JA, Kon SS, et al. Digital habits of PR service-users: implications for home-based interventions during the COVID-19 pandemic. Chron Respir Dis 2020 Jun 30;17:1479973120936685 [FREE Full text] [doi: 10.1177/1479973120936685] [Medline: 32602361]
42. Janaudis-Ferreira T, Tansey CM, Harrison SL, Beaurepaire CE, Goodridge D, Bourbeau J, et al. A qualitative study to inform a more acceptable pulmonary rehabilitation program after acute exacerbation of chronic obstructive pulmonary disease. Annals ATS 2019 Sep;16(9):1158-1164. [doi: 10.1513/annalsats.201812-854oc]

Abbreviations

AECOPD: acute exacerbation of chronic obstructive pulmonary disease
BCKQ: Bristol COPD Knowledge Questionnaire
COPD: chronic obstructive pulmonary disease
PR: pulmonary rehabilitation
RCT: randomized controlled trial
SPACE: Self-management Program of Activity Coping and Education

Edited by L Buis; submitted 23.06.20; peer-reviewed by A Agarwal, IV George, N Subedi; comments to author 30.08.20; revised version received 06.10.20; accepted 29.10.20; published 11.06.21.

Please cite as:
Houchen-Wolloff L, Orme M, Barradell A, Clinch L, Chaplin E, Gardiner N, Singh SJ
Web-Based Self-management Program (SPACE for COPD) for Individuals Hospitalized With an Acute Exacerbation of Chronic Obstructive Pulmonary Disease: Nonrandomized Feasibility Trial of Acceptability
JMIR Mhealth Uhealth 2021;9(6):e21728
URL: https://mhealth.jmir.org/2021/6/e21728
doi:10.2196/21728
PMID:34114960

Review

Efficacy of Mobile Health in Patients With Low Back Pain: Systematic Review and Meta-analysis of Randomized Controlled Trials

Mingrong Chen[1,2*], BS; Tingting Wu[1,2*], BS; Meina Lv[1,2], BS; Chunmei Chen[3*], PhD; Zongwei Fang[1,2], BS; Zhiwei Zeng[1,2], BS; Jiafen Qian[1,2], BS; Shaojun Jiang[1,2], BS; Wenjun Chen[1,2], BS; Jinhua Zhang[1,2], PhD

[1]Department of Pharmacy, Fujian Medical University Union Hospital, Fuzhou, China

[2]College of Pharmacy, Fujian Medical University, Fuzhou, China

[3]Department of Neurosurgery, Fujian Medical University Union Hospital, Fuzhou, China

[*]these authors contributed equally

Corresponding Author:
Jinhua Zhang, PhD
Department of Pharmacy
Fujian Medical University Union Hospital
#29 Xinquan Road
Fuzhou, 350001
China
Phone: 86 591 83357896 ext 8301
Fax: 86 591 83357896
Email: pollyzhang2006@126.com

Abstract

Background: Low back pain is one of the most common health problems and a main cause of disability, which imposes a great burden on patients. Mobile health (mHealth) affects many aspects of people's lives, and it has progressed rapidly, showing promise as an effective intervention for patients with low back pain. However, the efficacy of mHealth interventions for patients with low back pain remains unclear; thus, further exploration is necessary.

Objective: The purpose of this study was to evaluate the efficacy of mHealth interventions in patients with low back pain compared to usual care.

Methods: This was a systematic review and meta-analysis of randomized controlled trials designed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) statement standard. We searched for studies published in English before October 2020 in the PubMed, EMBASE, Web of Science, and Cochrane Library databases. Two researchers independently scanned the literature, extracted data, and assessed the methodological quality of the included studies. Bias risks were assessed using the Cochrane Collaboration tool. We used RevMan 5.4 software to perform the meta-analysis.

Results: A total of 9 studies with 792 participants met the inclusion criteria. The simultaneous use of mHealth and usual care showed a better reduction in pain intensity than usual care alone, as measured by the numeric rating scale (mean difference [MD] –0.85, 95% CI –1.29 to –0.40; *P*<.001), and larger efficacy in reducing disability, as measured by the Rolland-Morris Disability Questionnaire (MD –1.54, 95% CI –2.35 to –0.73; *P*<.001). Subgroup analyses showed that compared with usual care, mHealth using telephone calls significantly reduced pain intensity (MD –1.12, 95% CI –1.71 to –0.53; *P*<.001) and disability score (MD –1.68, 95% CI –2.74 to –0.63; *P*<.001). However, without the use of telephone calls, mHealth had no obvious advantage over usual care in improving pain intensity (MD –0.48, 95% CI –1.16 to 0.20; *P*=.16) and the disability score (MD –0.41, 95% CI –1.88 to 1.05; *P*=.58). The group that received a more sensitive feedback intervention showed a significantly reduced disability score (MD –4.30, 95% CI –6.95 to –1.69; *P*=.001).

Conclusions: The use of simultaneous mHealth and usual care interventions has better efficacy than usual care alone in reducing pain intensity and disability in patients with low back pain. Moreover, the results of subgroup analysis revealed that mHealth using telephone calls might play a positive role in improving pain intensity and disability in patients with low back pain.

(JMIR Mhealth Uhealth 2021;9(6):e26095) doi:10.2196/26095

KEYWORDS

mobile health; mHealth; low back pain; meta-analysis; pain intensity; disability

Introduction

Low back pain is one of the most common health problems worldwide, and is a main cause of disability according to the latest Global Burden of Disease Study [1]. In 2015, the global prevalence of low back pain with restricted mobility reached 7.3%, indicating that at any given time, low back pain affects 540 million people, representing a wide range of individuals. Low back pain occurs in high-, middle-, and low-income countries among all age groups, ranging from children to older adults [2]. Moreover, in industrialized countries, the lifetime prevalence of nonspecific low back pain is approximately 60%-70%. This not only affects people's private lives but also causes activity restrictions and work absences. At the same time, low back pain imposes a high economic burden on society [3]. Although physical exercise, conventional therapies, and cognitive behavioral therapy are the most effective nondrug conservative treatments to improve symptoms of low back pain, judging from the implementation of the traditional clinical model of management, the results obtained with these methods are not in line with expectations [4,5]. This finding may be due to the high degree of patient participation required, and the identification of issues and self-management that require patients to follow and complete time-intensive treatment plans independently at home [6]. These components may be the neglected aspects of usual care, which could include occupational therapy; physical therapy; or advice from a general practitioner (eg, family doctor), a specialist, or guidelines.

The term mobile health (mHealth) refers to the use of mobile devices such as mobile phones, patient monitoring equipment, and other wireless devices to provide medical support and health management [7], which may benefit health care providers by exerting positive effects on patient education, diagnosis, and management as components of the health delivery processes [8,9]. The devices for mHealth also include mobile phone software, text messaging, telephone calls, real-time monitoring (eg, motion sensor biofeedback), and network-based game consoles (eg, Nintendo Switch, Nintendo Wii). Compared with the traditional clinical model of management, mHealth can increase the feasibility and rationality of clinical treatment expectations by promoting patients' adherence to the treatment plan. mHealth can also provide a basis for formulating treatment plans and compensate for the traditional model's shortcomings to a certain extent [10]. In addition, mHealth can help achieve universal health service coverage by overcoming geographical barriers, thereby increasing the number of paths by which patients can enter medical care systems, and providing medical care services to people in remote areas and communities with insufficient services and inadequate conditions. As no additional medical equipment or time utilization is needed to use mHealth interventions, the expenditures for clinical data monitoring and educational/information exchange between doctors and patients are lower than those of face-to-face services [11]. Moreover, mHealth has been used in many aspects of people's lives to help them adapt to various health conditions and problems, including those related to mental health [12,13], heart failure [14], and smoking cessation [15,16].

At present, few meta-analyses have been performed on the efficacy of mHealth for patients with low back pain, and therefore the ability to provide more effective or accurate clinical advice is limited. In light of the various advantages of mHealth that are different from usual care and the current global status of low back pain, we performed a meta-analysis to clarify the efficacy of mHealth for patients with low back pain.

Methods

Search Strategy

Data were retrieved from the PubMed, Embase, Web of Science, and Cochrane Library databases. We searched for studies in English published until October 2020. The key search strings consisted of two concepts: mHealth and low back pain (Textbox 1). The reference lists of the retrieved studies were checked for identifying further relevant studies.

Textbox 1. Example of the search strategy (EMBASE).

- Search 1: ("mobile application" OR "telemedicine" OR "text messaging" OR "mobile phone" OR "smartphone" OR "social media" OR "internet")/exp
- Search 2: (mobile OR "portable software application" OR tele* OR mhealth OR ehealth OR "e health" OR "mhealth" OR ?phone* OR text* OR "short message" OR sms OR app OR apps OR digital* OR web* OR internet* OR ?media OR wireless OR computer OR video* OR bluetooth OR blog* OR online OR electronic OR "mp3 player" OR "mp4 player" OR wechat OR whatsapp OR twitter OR "virtual reality" OR "interactive voice response" OR facebook OR networking): title/abstract/keywords
- Search 3: Search 1 OR Search 2
- Search 4: "low back pain"/exp OR "backache"/exp
- Search 5: ("low back pain" OR "low back ache" OR "low backache" OR "lower back pain" OR "back disorder" OR backache OR "back pain" OR lumbago OR dorsalgia OR coccyx OR coccydynia OR sciatica OR ischialgia OR spondylosis): title/abstract/keywords
- Search 6: Search 4 OR Search 5
- Search 7: Search 3 AND Search 6

Study Inclusion Criteria

The following inclusion criteria were used: (1) the study design was a randomized controlled trial (RCT); (2) mHealth (eg, mobile phone, computer, motion sensor biofeedback, and network-based game consoles) and usual care (eg, exercise and/or advice) were used simultaneously in the experimental group, and usual care or usual care and placebo were used in the control group; (3) participants were confirmed to have low back pain; and (4) the outcomes were measured using the numeric rating scale (NRS) and/or Roland-Morris Disability Questionnaire (RMDQ), with data expressed as mean (SD). Two researchers selected the studies independently in accordance with the above criteria.

Study Exclusion Criteria

Studies were excluded from the meta-analysis if: (1) not all of the participants were diagnosed with low back pain; (2) the study population included pregnant women or patients recovering after spinal surgery; or (3) email had been used as an intervention for office workers in the workplace. The latter criterion was based on previous studies [17-20] indicating a consistently high degree of patient compliance. We believe that the specific combination of being an office worker and receiving email at the workplace conveys a highly sensitive nature of the feedback, resulting in high patient compliance. This finding is inconsistent with the general usage of email in mHealth and is not universal. Therefore, this meta-analysis did not include interventions for office workers using email in the workplace.

Data Extraction

The required data were extracted independently by two researchers and crosschecked to avoid potential data extraction errors. Disagreements during the extraction process were resolved by seeking the opinion of a third researcher. The extracted information included the first author's name, year of publication, sample size, age, sex ratio, and participants' scores on the NRS and RMDQ. The final postintervention data with the longest follow-up time was used in our analysis if the study (ie, [21-23]) reported results from a different period.

Data Analysis

We assigned participants who used telephone calls, internet/email, mobile phones, or other mHealth methods and usual care at the same time to the mHealth experimental group, and those who used usual care alone to the control group. RevMan 5.4 software was used for the meta-analysis, with the mean difference (MD), SD, and 95% CI as the statistics of interest. The overall pooled effect estimate was assessed using Z - statistics, and statistical significance was considered at $P<.05$. The χ^2 test was used to examine the heterogeneity of the results: if $P\geq.10$ and $I^2\leq 50\%$, a fixed-effects model was used for the meta-analysis; otherwise, a random-effects model was used.

Given reports of positive effects of telephone calls on patients' self-management and compliance [24], we defined subgroups based on two indicators: whether the intervention used telephone calls or more sensitive feedback methods (ie, motion sensor biofeedback). We examined the effect of telephone calls on the following outcomes: (1) whether the use of telephone calls affects NRS scores (the experimental group using mHealth was divided into two groups based on use of telephone calls and no telephone calls); and (2) whether the use of telephone calls or more sensitive feedback interventions affects RMDQ scores (the experimental group using mHealth was divided into three groups based on use of telephone calls, no telephone calls, and use of more sensitive feedback interventions such as motion sensor biofeedback).

Quality Assessment

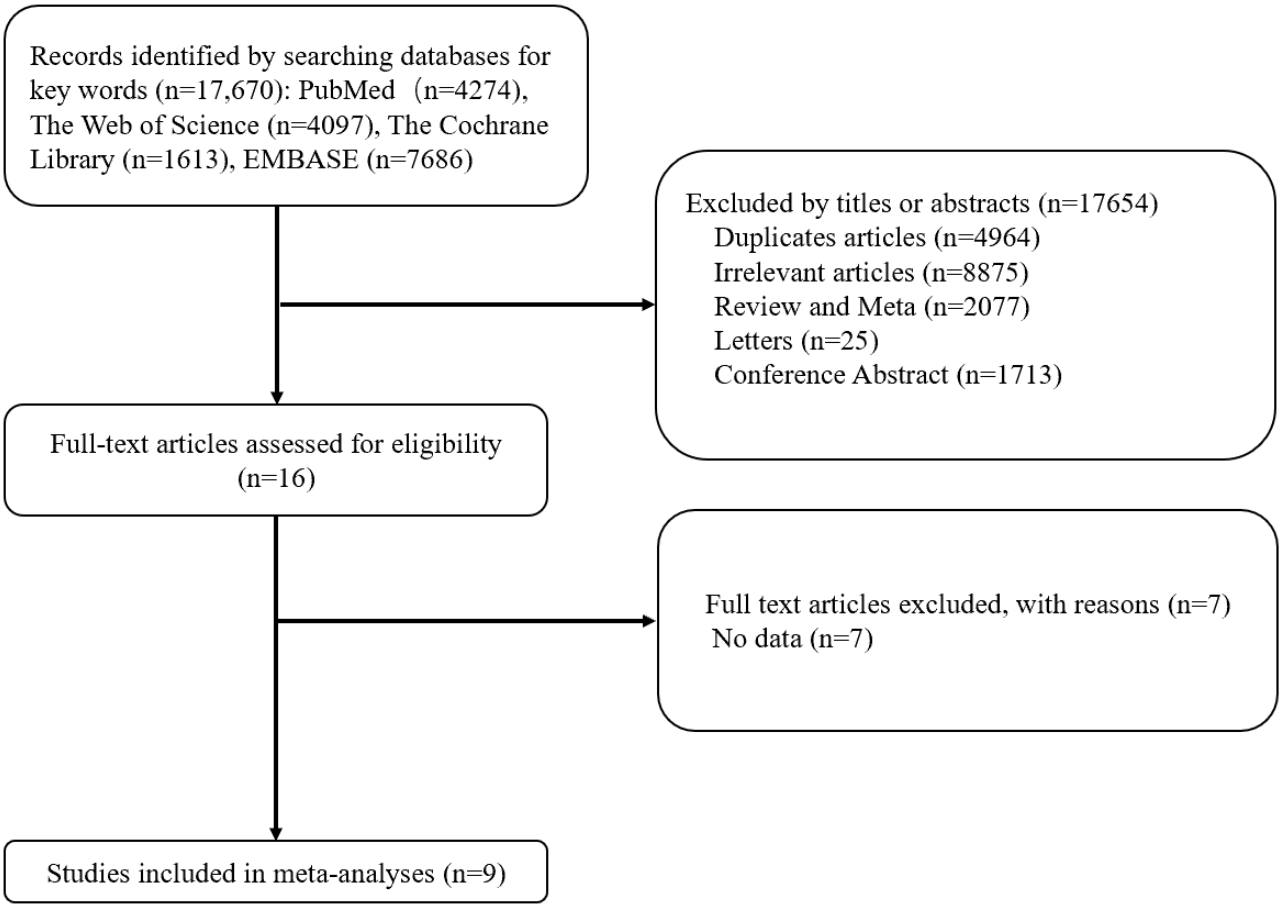
The risk of bias was assessed using the Cochrane Collaboration Tool for Assessing Risk of Bias in Randomised Trials [25]. We evaluated seven aspects of the studies: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessments, incomplete outcome data, selective reporting, and other biases. Risk of bias ratings of "low," "high," and "unclear" were assigned to each aspect of each study. Two researchers performed the quality assessments of the included studies. Differences in opinions of the two researchers were resolved through discussion and decision. A third researcher was consulted for disagreements, and a decision was made after a discussion.

Results

Study Selection

A total of 17,670 studies were identified during the initial examination. After all studies were screened and filtered using the inclusion and exclusion criteria, 9 studies were included in the final meta-analysis. These studies comprised 792 participants, 407 of whom were allocated to the mHealth group and 385 were allocated to the usual care group, as shown in the flowchart in Figure 1.

Figure 1. Flowchart of the screening and selection of studies.



Study Characteristics

Features of the included studies and outcome data are summarized in Multimedia Appendix 1 and Table 1, respectively. Of the 9 included studies, 3 from research institutions were performed in Australia [21,22,26]; 1 was performed in the United States [23]; and 1 study each was performed in India, China, Brazil, the United Kingdom, and Italy. The proportion of female participants ranged from 50% to 74%, and the mean age of the participants ranged from 40 to 68 years. The follow-up period of the included studies ranged from 4 weeks to 12 months.

Table 1. Pain and disability scores at baseline and after the interventions.

Study	Pain intensity: NRS[a] (score range 0-10), mean (SD)				Disability: RMDQ[b] (score range 0-24), mean (SD)			
	mHealth[c] group		Control group		mHealth group		Control group	
	Baseline	After intervention	Baseline	After intervention	Baseline	After intervention	Baseline	After intervention
Amorim et al [26]	5.3 (1.9)	3.8 (2.4)	5.1 (1.4)	4.0 (3.4)	8.9 (5.4)	5.7 (5.3)	9.0 (6.1)	6.0 (5.7)
Bernardelli et al [27]	NA[d]	NA	NA	NA	6.3 (4.4)	3.8 (3.9)	6.4 (4.9)	4.3 (4.2)
Chhabra et al [28]	7.3 (1.9	3.3 (1.7)	6.6 (2.1)	3.2 (2.7)	NA	NA	NA	NA
Damush et al [23]	NA	NA	NA	NA	14.7 (6.7)	9.1 (6.8)	13.9 (6.8)	11.3 (8.1)
Geraghty et al [29]	Ⅰ: 4.0 (2.6); Ⅱ:4.5 (2.6)	Ⅰ: 3.6 (2.5); Ⅱ: 3.1 (2.3)	3.6 (3.1)	4.0 (2.5)	Ⅰ: 6.6 (4.6); Ⅱ: 7.7 (4.7)	Ⅰ: 5.8 (4.5); Ⅱ: 5.1 (5.1)	6.8 (4.9)	6.3 (5.1)
Kent et al [21]	NA	NA	NA	NA	11.8 (8.8)	7.2 (2.6)	11.3 (7.0)	11.0 (1.3)
Monteiro-Junior et al [30]	6.5 (1.1)	1.7 (1.9)	6.6 (1.2)	1.4 (2.9)	NA	NA	NA	NA
Petrozzi et al [22]	5.1 (1.8)	3.0 (2.1)	4.9 (2.0)	4.0 (2.1)	9.9 (4.2)	4.2 (3.7)	9.9 (4.7)	5.3 (5.1)
Yang et al [31]	NA	NA	NA	NA	6.00 (3.74)	4.40 (3.05)	12.00 (3.61)	11.70 (5.69)

[a]NRS: numeric rating scale.

[b]RMDQ: Rolland-Morris Disability Questionnaire.

[c] mHealth: mobile health.

[d]NA: not applicable (not assessed).

Risk of Bias in the Included Studies

The Cochrane Collaboration Risk of Bias Tool was used to assess the risk of bias of the 9 included studies. Seven studies used computer-generated random numbers [21,22,26-30] and 2 studies did not indicate use of a sequence generation method [23,31]. Two studies did not report allocation concealment [23,30]; therefore, the risk of selection bias was relatively low. The risk of performance bias was found to be relatively high. Participants were blinded to the treatment conditions in 4 of the studies [21,23,29,30], 2 studies lacked sufficient information about whether the participants were blinded [26,27], and participants were not blinded in 3 studies [22,28,31]. Investigators were blinded to the outcomes in all of the included studies. An unclear risk of attrition bias was found in 1 study [31] and a high risk was found in 2 studies, as the dropout rate was relatively high [26,29]. The risk for reporting bias was low in all studies, and the risk for other types of bias was high in 1 study [28].

The overall risk of bias was relatively low, but performance bias was relatively high, as 3 of the studies used a single-blinded method [22,28,31] (Figure 2). However, it should be noted that the effects of blinding on the results of studies in the fields of rehabilitation and physical therapy are currently unclear [32,33].

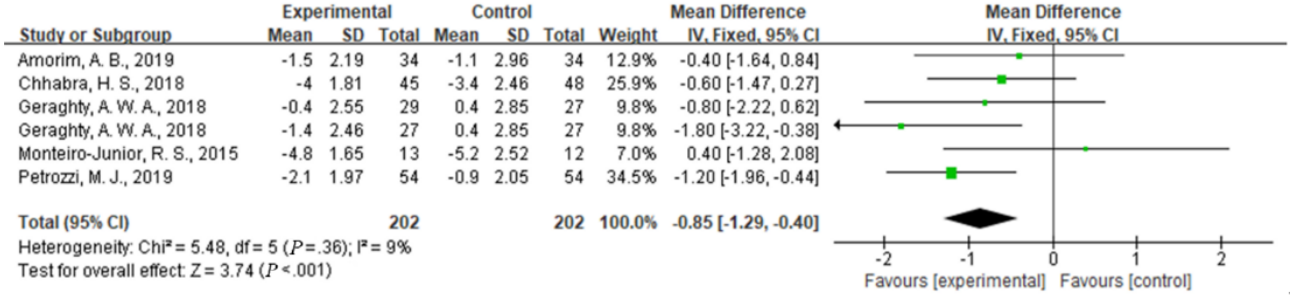
Figure 2. Risk assessment of bias in the included studies.

Meta-analysis

Comparison of Pain Intensity

Compared with usual care, the simultaneous interventions of mHealth and usual care were more effective in reducing pain, as indicated by the NRS scores of 404 participants in 5 studies (MD –0.85, 95% CI –1.29 to –0.40; I^2=9%; *P*<.001) (Figure 3).

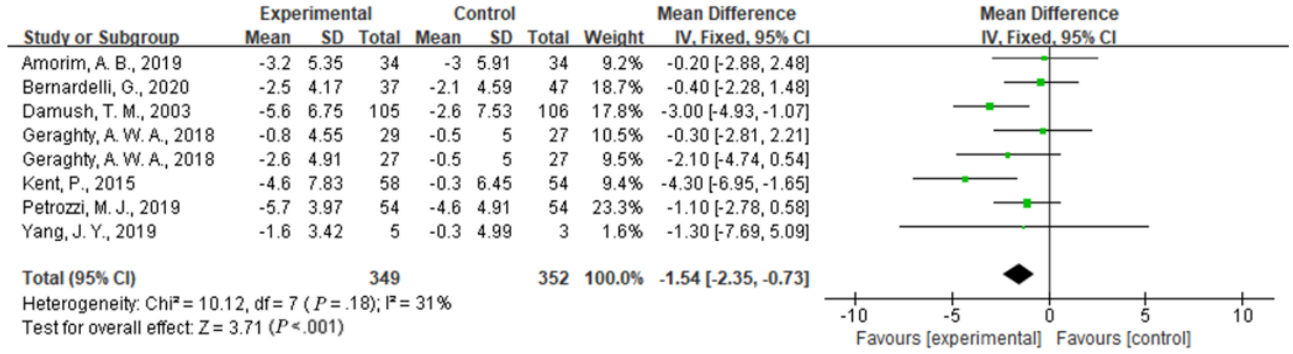
Figure 3. Forest plot of the efficacy of mobile health and traditional health interventions in reducing pain intensity.



Comparison of Disability

Compared with usual care, the simultaneous interventions of mHealth and usual care had a larger effect on reducing disability in patients with low back pain, as indicated by the RMDQ scores of 885 participants in 8 studies (MD –1.54, 95% CI –2.35 to –0.73; I^2 =31%; *P*<.001) (Figure 4).

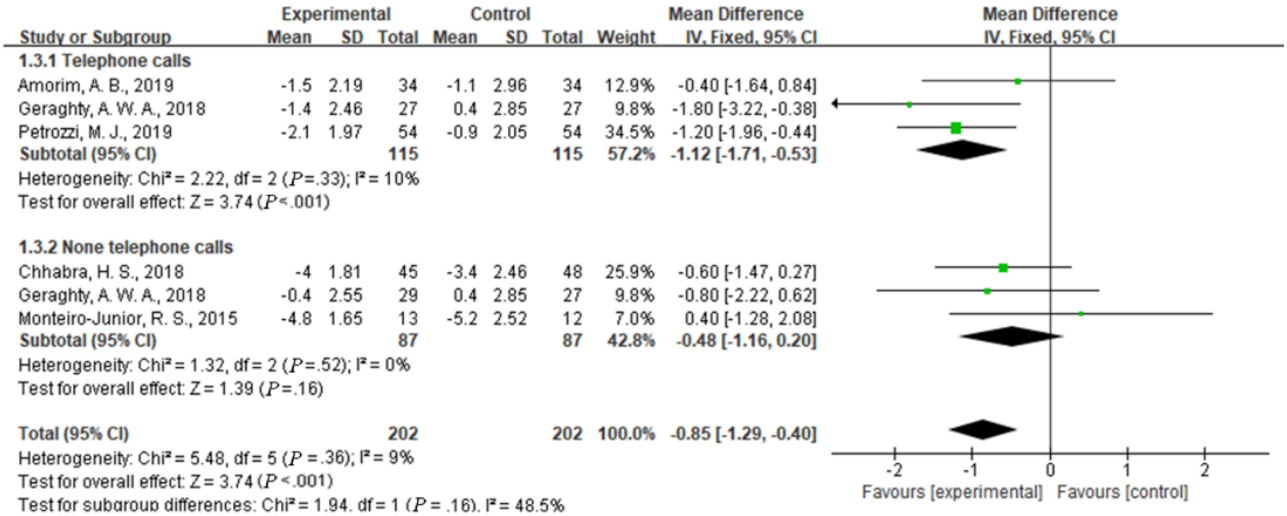
Figure 4. Forest plot of the efficacy of mobile health and traditional health interventions on disability.



Subgroup Analysis

We evaluated pain intensity, as measured by the NRS, to examine the efficacy of mHealth using telephone calls. We performed a subgroup analysis of participants in the mHealth experimental group that used telephone calls and those who were in an intervention group that did not use telephone calls. A difference was found between the telephone group and nontelephone group, although it was not statistically significant (I^2=48.5%, *P*=.16). Compared with usual care, mHealth using telephone calls significantly reduced pain intensity in 3 studies (MD –1.12, 95% CI –1.71 to –0.53; I^2 =10%, *P*<.001); however, without the use of telephone calls, mHealth had no obvious advantage over usual care (MD –0.48, 95% CI –1.16 to 0.20; I^2 0%, *P*=.16) (Figure 5).

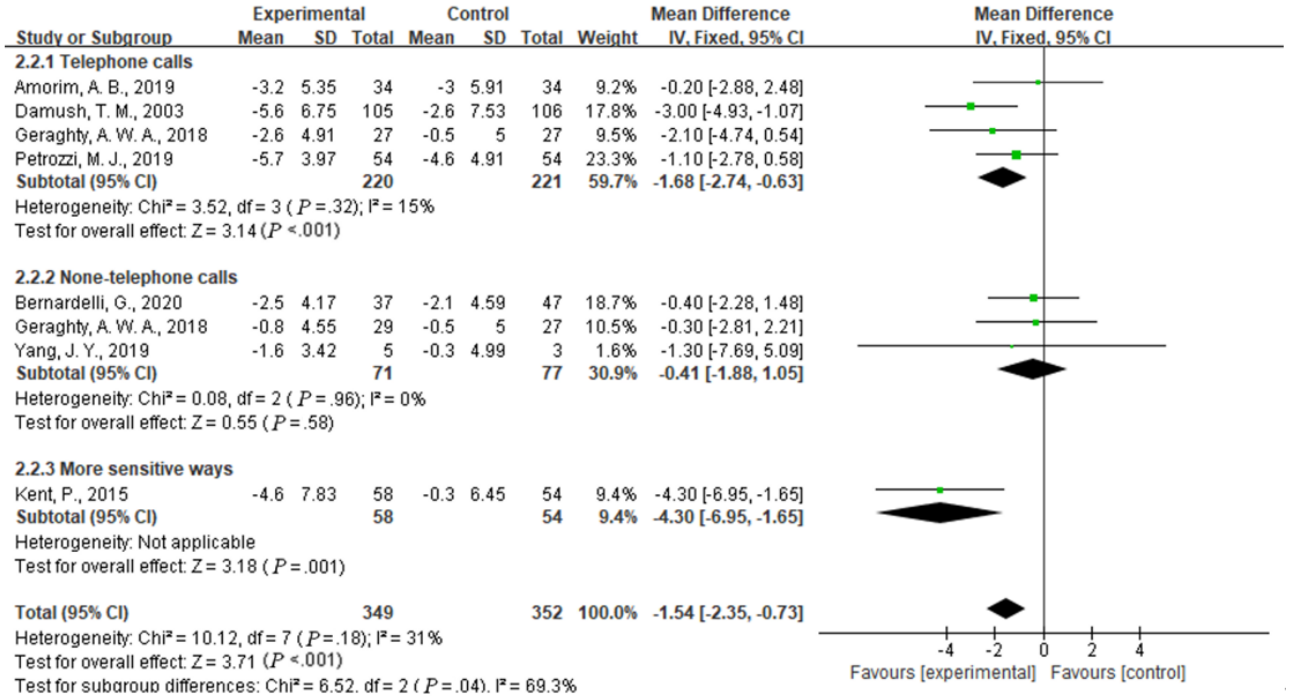
Figure 5. Efficacy of telephone calls use in reducing pain intensity.



We evaluated disability, as measured by the RMDQ, to examine the efficacy of different types of mHealth. We performed a subgroup analysis of participants in the mHealth experimental group that used telephone calls, did not use telephone calls, or used a more sensitive feedback intervention. The analysis indicated a significant difference between the telephone calls

group, the more sensitive feedback intervention group, and the nontelephone group (I^2=69.3%, *P*=.04). Compared with the group that received usual care, the experimental mHealth group that involved telephone calls showed a significantly reduced disability score in 4 studies (MD –1.68, 95% CI –2.74 to –0.63; I^2 = 15%, *P*<.001). The group that received a more sensitive feedback intervention showed a significantly reduced disability score in 1 study (MD –4.30, 95% CI –6.95 to –1.69; *P*=.001), and the group that did not receive an intervention with telephone calls showed no significant difference in their RMDQ scores from the two other groups in 3 studies (MD –0.41, 95% CI –1.88 to 1.05; I^2=0%, *P*=.58) (Figure 6).

Figure 6. Efficacy of telephone use and the more sensitive feedback intervention in reducing disability.



Publication Bias and Sensitivity Analysis

A sensitivity analysis was performed using one-by-one elimination of studies that reported the outcomes of the NRS and the RMDQ. No significant change was found in the outcomes, indicating that the results were stable. According to the Cochrane Group, a funnel plot to detect publication bias is not recommended when fewer than 9 studies are included in a meta-analysis [34]. Hence, a funnel plot was not used to detect publication bias because of the small number of studies in our meta-analysis.

Discussion

Principal Findings

This meta-analysis of 9 studies with 792 patients revealed a significant positive effect of the simultaneous interventions of mHealth and usual care compared with usual care only on patients with low back pain. The mHealth intervention performed significantly better than usual care on measures of pain intensity (MD –0.85, 95% CI –1.29 to –0.40; I^2=9%, *P*<.001) and disability (MD –1.54, 95% CI –2.35 to –0.73; I^2=31%, *P*<.001). The subgroup analysis of the scores on the NRS and RMDQ showed that the use of telephone calls or more sensitive feedback devices for the intervention might be superior to other types of mHealth interventions or usual care in terms of improving pain intensity and disability. This conclusion is consistent with a study by Niznik et al [24] who reported advantages of telephone calls, and concluded that telephone calls have positive effects on clinical disease management, patient management, and patient compliance. This study has similarities and differences with the results of another study [35] that examined differences in mHealth between a web-based health program and an mHealth-based program. No significant effects on pain or disability were found among participants in the web health program versus the controls. Compared with the controls, the participants in the trials on mHealth-based programs reported clinically significant effects on pain intensity and disability.

Our study demonstrates the importance of telephone calls in mHealth. We believe that telephone calls may be one of the main effective types of mHealth with great positive effects on patients in reducing pain and disability. As one of the mobile medical methods, telephone calls might be superior to other types of mHealth for the following reasons. First, according to Simblett et al [36], a major challenge of mHealth is the high dropout rate of participants with sensors and the usability of apps. Therefore, in contrast to other types of mHealth, we believe that active telephone calling from the treatment unit instead of the passive use of software and websites can greatly improve the enthusiasm and compliance of patients, thereby ensuring a positive impact of mHealth on patients. Another option is to use motion sensor biofeedback to achieve real-time communication and feedback to improve efficacy, which is consistent with the opinion of Sim [37]. Second, mHealth has potential to facilitate the achievement of universal coverage for

health services by overcoming geographical barriers, increasing the number of pathways to medical care, and providing medical care services to people in remote areas and communities with insufficient services and inadequate conditions [11]. However, in most countries, especially in remote areas, the network infrastructure is far less stable than telephone calls. The costs in money and time of developing software or websites in business or medical settings are much higher than those associated with telephone calls.

To our knowledge, subgroup analyses of the efficacy of telephone calls have not been performed to date. Since few studies have examined the impact of mHealth on patients with low back pain, there are no related articles for comparison. Only a systematic review and meta-analysis of 5 articles performed by Du et al [35] on the effectiveness of mHealth in the self-management of patients with chronic low back pain was found in the published literature. The results revealed that mHealth-based self-management might play a positive role in improving short-term pain intensity and short-term disability in patients with chronic low back pain. After careful reading of this meta-analysis, we found similarities and differences with our study. The endpoint data of the included RCTs were also extracted by Du et al [35] for the meta-analysis. However, in contrast to their study, we included more articles in our meta-analysis, and calculated differences between the baseline and endpoint data of each included RCT. We believe that this method is more accurate and that it further supports our conclusion. Most previous meta-analyses related to mHealth did not distinguish between the use of mHealth alone and the simultaneous use of mHealth and routine care, nor did they restrict the intervention methods of the control group. We chose usual care and mHealth for the intervention group, and usual care only for the control group, and we believe that such a comparison yielded conclusions that are more reliable than other comparisons.

Establishing clinical relevance is the key to whether mHealth can be used in patients with low back pain. Yet, small effects (–0.85) are observed at the group level for pain intensity when compared to the control group, which do not meet the minimal clinically important difference criterion of –1.77 [38]. However, as one of the intervention methods used simultaneously with usual care, mHealth can significantly improve the curative effect in reducing pain intensity and disability in patients with low back pain, while reducing human resources and time costs. Therefore, this method is worthy of adoption.

The objective of this study was to examine the influence of mHealth interventions on the pain intensity and disability of patients with low back pain. Our investigation highlights differences between the intervention of usual care alone and the simultaneous use of usual care and mHealth. Compared with using usual care alone, the intervention of telephone calls had a significant beneficial effect on patients' disability. These findings are expected to provide guidance for clinical decisions and contribute to this field.

Limitations

Our study has several limitations. First, this meta-analysis may be biased if the literature search failed to identify all trials reporting on differences between mHealth and usual care or if the selection criteria for including trials were applied in a subjective manner. To reduce these risks, we performed thorough searches across multiple literature databases and clinical trial databases, and used explicit criteria for study selection and data extraction and analysis. Second, mHealth may have specific effects that vary by the type of low back pain. That is, to better evaluate the efficacy, and save human resources and time costs, passive sensing in mHealth may be more suitable for chronic low back pain, whereas active sensing may be more suitable for acute low back pain, which can be administered multiple times a day to capture short-term variations in responses [37]. However, owing to the insufficient number of studies on acute and subacute low back pain, we were unable to perform a subgroup analysis according to the type of back pain, and therefore this issue should be examined in the future. Finally, as this was a study-level rather than participant-level meta-analysis, we were able to analyze univariate associations, but not multivariate associations of baseline features with outcomes.

Conclusion

The results of this meta-analysis suggest that the simultaneous interventions of mHealth and usual care, compared with usual care alone, are significantly better for reducing pain intensity and disability in patients with low back pain. The use of telephone calls or more sensitive feedback interventions may further increase the positive effects of these simultaneous interventions on the disability of patients with low back pain. The wider use of mHealth may contribute significantly to the population of patients with low back pain. Therefore, the simultaneous interventions of mHealth and usual care may be a promising method worth considering.

Acknowledgments

This study was funded by the Joint Funds for the Innovation of Science and Technology, Fujian Province (grant number 2018Y9060), the Natural Science Foundation of Fujian Province, China (2018Y0037), and the Fujian Provincial Health Technology Project, China (2019-CX-19).

Authors' Contributions

JZ initiated the study. ZF and JQ selected the studies for inclusion in the meta-analysis. SJ and ZZ performed the quality assessments of the included studies. TW and ML performed the data extraction and analyses, and MC and CC drafted the first version of the

manuscript. WC advised on data analysis. JZ, TW, and MC critically reviewed and revised the manuscript. All authors made a substantial contribution to the concept and design of the study, interpretation of the data, and review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the included studies.
[DOCX File , 19 KB - mhealth_v9i6e26095_app1.docx]

References

1. GBD 2016 Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
2. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet 2018 Jun 09;391(10137):2356-2367. [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
3. Hoy D, Brooks P, Blyth F, Buchbinder R. The epidemiology of low back pain. Best Pract Res Clin Rheumatol 2010 Dec;24(6):769-781. [doi: 10.1016/j.berh.2010.10.002] [Medline: 21665125]
4. Chou R, Côté P, Randhawa K, Torres P, Yu H, Nordin M, et al. The Global Spine Care Initiative: applying evidence-based guidelines on the non-invasive management of back and neck pain to low- and middle-income communities. Eur Spine J 2018 Sep;27(Suppl 6):851-860. [doi: 10.1007/s00586-017-5433-8] [Medline: 29460009]
5. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 04;166(7):514-530 [FREE Full text] [doi: 10.7326/M16-2367] [Medline: 28192789]
6. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med 2012 May;27(5):520-526 [FREE Full text] [doi: 10.1007/s11606-011-1931-2] [Medline: 22127797]
7. Global Diffusion of eHealth: Making Universal Health Coverage Achievable. Geneva: World Health Organization; 2017.
8. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
9. Mosa ASM, 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] [Medline: 22781312]
10. 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] [Medline: 25803266]
11. Wheeler TS, Michael Vallis T, Giacomantonio NB, Abidi SR. Feasibility and usability of an ontology-based mobile intervention for patients with hypertension. Int J Med Inform 2018 Nov;119:8-16. [doi: 10.1016/j.ijmedinf.2018.08.002] [Medline: 30342690]
12. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
13. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
14. 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] [Medline: 27302310]
15. BinDhim NF, McGeechan K, Trevena L. Smartphone Smoking Cessation Application (SSC App) trial: a multicountry double-blind automated randomised controlled trial of a smoking cessation decision-aid 'app'. BMJ Open 2018 Jan 21;8(1):e017105 [FREE Full text] [doi: 10.1136/bmjopen-2017-017105] [Medline: 29358418]
16. Brinker TJ, Brieske CM, Esser S, Klode J, Mons U, Batra A, et al. A face-aging app for smoking cessation in a waiting room setting: pilot study in an HIV outpatient clinic. J Med Internet Res 2018 Aug 15;20(8):e10976 [FREE Full text] [doi: 10.2196/10976] [Medline: 30111525]

17. Franklin PD, Rosenbaum PF, Carey MP, Roizen MF. Using sequential e-mail messages to promote health behaviors: evidence of feasibility and reach in a worksite sample. J Med Internet Res 2006 Mar 30;8(1):e3 [FREE Full text] [doi: 10.2196/jmir.8.1.e3] [Medline: 16585028]
18. Plotnikoff RC, McCargar LJ, Wilson PM, Loucaides CA. Efficacy of an E-mail intervention for the promotion of physical activity and nutrition behavior in the workplace context. Am J Health Promot 2005;19(6):422-429. [doi: 10.4278/0890-1171-19.6.422] [Medline: 16022206]
19. Del Pozo-Cruz B, Adsuar JC, Parraca J, Del Pozo-Cruz J, Moreno A, Gusi N. A web-based intervention to improve and prevent low back pain among office workers: a randomized controlled trial. J Orthop Sports Phys Ther 2012 Oct;42(10):831-841. [doi: 10.2519/jospt.2012.3980] [Medline: 22951407]
20. Yap TL, Davis LS, Gates DM, Hemmings AB, Pan W. The effect of tailored E-mails in the workplace. Part I. Stage movement toward increased physical activity levels. AAOHN J 2009 Jul;57(7):267-273. [Medline: 19639858]
21. Kent P, Laird R, Haines T. The effect of changing movement and posture using motion-sensor biofeedback, versus guidelines-based care, on the clinical outcomes of people with sub-acute or chronic low back pain-a multicentre, cluster-randomised, placebo-controlled, pilot trial. BMC Musculoskelet Disord 2015 May 29;16:131 [FREE Full text] [doi: 10.1186/s12891-015-0591-5] [Medline: 26022102]
22. Petrozzi MJ, Leaver A, Ferreira PH, Rubinstein SM, Jones MK, Mackey MG. Addition of MoodGYM to physical treatments for chronic low back pain: A randomized controlled trial. Chiropr Man Therap 2019;27:54 [FREE Full text] [doi: 10.1186/s12998-019-0277-4] [Medline: 31673330]
23. Damush TM, Weinberger M, Perkins SM, Rao JK, Tierney WM, Qi R, et al. The long-term effects of a self-management program for inner-city primary care patients with acute low back pain. Arch Intern Med 2003 Nov 24;163(21):2632-2638. [doi: 10.1001/archinte.163.21.2632] [Medline: 14638564]
24. Niznik JD, He H, Kane-Gill SL. Impact of clinical pharmacist services delivered via telemedicine in the outpatient or ambulatory care setting: A systematic review. Res Social Adm Pharm 2018 Aug;14(8):707-717. [doi: 10.1016/j.sapharm.2017.10.011] [Medline: 29100941]
25. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
26. Amorim AB, Pappas E, Simic M, Ferreira ML, Jennings M, Tiedemann A, et al. Integrating mobile-health, health coaching, and physical activity to reduce the burden of chronic low back pain trial (IMPACT): a pilot randomised controlled trial. BMC Musculoskelet Disord 2019 Feb 11;20(1):71 [FREE Full text] [doi: 10.1186/s12891-019-2454-y] [Medline: 30744606]
27. Bernardelli G, Vigna L, Nava C, De Gennaro Colonna V, Andersen LL, Consonni D, et al. Physical activity in healthcare workers with low back pain: effects of the Back-FIT randomized trial. J Occup Environ Med 2020 Jun;62(6):e245-e249. [doi: 10.1097/JOM.0000000000001844] [Medline: 32097286]
28. 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] [Medline: 30324496]
29. Geraghty AWA, Stanford R, Stuart B, Little P, Roberts LC, Foster NE, et al. Using an internet intervention to support self-management of low back pain in primary care: findings from a randomised controlled feasibility trial (SupportBack). BMJ Open 2018 Mar 09;8(3):e016768 [FREE Full text] [doi: 10.1136/bmjopen-2017-016768] [Medline: 29525768]
30. Monteiro-Junior RS, de Souza CP, Lattari E, Rocha NBF, Mura G, Machado S, et al. Wii-workouts on chronic pain, physical capabilities and mood of older women: a randomized controlled double blind trial. CNS Neurol Disord Drug Targets 2015;14(9):1157-1164. [doi: 10.2174/1871527315666151111120131] [Medline: 26556092]
31. Yang J, Wei Q, Ge Y, Meng L, Zhao M. Smartphone-based remote self-management of chronic low back pain: a preliminary study. J Healthc Eng 2019;2019:4632946. [doi: 10.1155/2019/4632946] [Medline: 30881606]
32. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. Am J Phys Med Rehabil 2017 Jan;96(1):34-44. [doi: 10.1097/PHM.0000000000000521] [Medline: 27149591]
33. Armijo-Olivo S, Dennett L, Arienti C, Dahchi M, Arokoski J, Heinemann AW, et al. Blinding in rehabilitation research: empirical evidence on the association between blinding and treatment effect estimates. Am J Phys Med Rehabil 2020 Mar;99(3):198-209. [doi: 10.1097/PHM.0000000000001377] [Medline: 31913147]
34. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010 Sep;25(9):603-605. [doi: 10.1007/s10654-010-9491-z] [Medline: 20652370]
35. Du S, Liu W, Cai S, Hu Y, Dong J. The efficacy of e-health in the self-management of chronic low back pain: A meta analysis. Int J Nurs Stud 2020 Jun;106:103507. [doi: 10.1016/j.ijnurstu.2019.103507] [Medline: 32320936]
36. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
37. Sim I. Mobile devices and health. N Engl J Med 2019 Sep 05;381(10):956-968. [doi: 10.1056/NEJMra1806949] [Medline: 31483966]

38. Haase I, Kladny B. Clinical relevance of changes in pain intensity in patients with specific back pain. Z Orthop Unfall 2021 Jan 18:online ahead of print. [doi: 10.1055/a-1304-3677] [Medline: 33461231]

Abbreviations

MD: mean difference
mHealth: mobile health
NRS: numeric rating scale
RCT: randomized controlled trial
RMDS: Roland-Morris Disability Questionnaire

Edited by L Buis; submitted 27.11.20; peer-reviewed by J Lurie, S Lilje; comments to author 25.01.21; revised version received 26.02.21; accepted 18.03.21; published 11.06.21.

Please cite as:
Chen M, Wu T, Lv M, Chen C, Fang Z, Zeng Z, Qian J, Jiang S, Chen W, Zhang J
Efficacy of Mobile Health in Patients With Low Back Pain: Systematic Review and Meta-analysis of Randomized Controlled Trials
JMIR Mhealth Uhealth 2021;9(6):e26095
URL: https://mhealth.jmir.org/2021/6/e26095
doi:10.2196/26095
PMID:34114965

Original Paper

An Intervention to Improve Medication Adherence in People With Heart Disease (Text4HeartII): Randomized Controlled Trial

Ralph Maddison[1], BHSc, MSc, PhD; Yannan Jiang[2], PhD; Ralph Stewart[3], PhD, MBChB; Tony Scott[4], MBChB; Andrew Kerr[5], MBChB, MD; Robyn Whittaker[6,7], PhD, MBChB; Jocelyn Benatar[3], MBChB; Anna Rolleston[8], PhD; Paul Estabrooks[9], PhD; Leila Dale[6], PhD

[1]Institute for Physical Activity and Nutrition, Deakin University, Burwood, Australia
[2]Department of Statistics, The University of Auckland, Auckland, New Zealand
[3]Department of Cardiology, Auckland District Health Board, Auckland, New Zealand
[4]Department of Cardiology, Waitemata District Health Board, Auckland, New Zealand
[5]Section of Epidemiology and Biostatistics and School of Medicine, University of Auckland, Auckland, New Zealand
[6]National Institute for Health Innovation, University of Auckland, Auckland, New Zealand
[7]Waitemata District Health Board, Auckland, New Zealand
[8]The Centre for Health, Tauranga, New Zealand
[9]Department of Health Promotion, Social and Behavioral Health, University of Nebraska Medical Centre, Nebraska, NE, United States

Corresponding Author:
Ralph Maddison, BHSc, MSc, PhD
Institute for Physical Activity and Nutrition
Deakin University
221 Burwood Highway
Burwood, 3125
Australia
Phone: 61 3 924 46218
Email: ralph.maddison@deakin.edu.au

Abstract

Background: Mobile health technologies have the potential to improve the reach and delivery of interventions for promoting long-term secondary prevention of coronary heart disease.

Objective: This study aims to determine the effectiveness of an SMS text messaging intervention (Text4HeartII) for improving adherence to medication and lifestyle changes over and above usual care in people with coronary heart disease at 24 and 52 weeks.

Methods: A two-arm, parallel, randomized controlled trial was conducted in New Zealand. Participants with a recent acute coronary syndrome were randomized to receive usual cardiac services alone (control, n=153) or a 24-week SMS text message program for supporting self-management plus usual cardiac services (n=153). The primary outcome was adherence to medication at 24 weeks, defined as a medication possession ratio of 80% or more for aspirin, statin, and antihypertensive therapy. Secondary outcomes included medication possession ratio at 52 weeks, self-reported medication adherence, adherence to healthy lifestyle behaviors, and health-related quality of life at 24 and 52 weeks.

Results: Participants were predominantly male (113/306, 80.3%) and European New Zealanders (210/306, 68.6%), with a mean age of 61 years (SD 11 years). Groups were comparable at baseline. National hospitalization and pharmacy dispensing records were available for all participants; 92% (282/306, 92.1%) of participants completed a 24-week questionnaire and 95.1% (291/306) of participants completed a 52-week questionnaire. Adherence with 3 medication classes were lower in the intervention group than in the control group (87/153, 56.8% vs 105/153, 68.6%, odds ratio 0.60, 95% CI 0.38-0.96; *P*=.03) and 52 weeks (104/153, 67.9% vs 83/153, 54.2%; odds ratio 0.56, 95% CI 0.35-0.89; *P*=.01). Self-reported medication adherence scores showed the same trend at 52 weeks (mean difference 0.3; 95% CI 0.01-0.59; *P*=.04). Moreover, self-reported adherence to health-related behaviors was similar between groups.

Conclusions: Text4HeartII did not improve dispensed medication or adherence to a favorable lifestyle over and above usual care. This finding contrasts with previous studies and highlights that the benefits of text interventions may depend on the context in which they are used.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12616000422426; http://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=370398.

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-018-2468-z

(JMIR Mhealth Uhealth 2021;9(6):e24952) doi:10.2196/24952

KEYWORDS

cardiovascular disease; self-management; text messaging; risk factors

Introduction

Background

Coronary heart disease (CHD) is a leading cause of premature death and disability worldwide [1]. Improved diagnosis, treatment, and management have substantially reduced the mortality rate of individuals living with CHD [2,3]. Following a cardiac event, clinical guidelines recommend people should participate in cardiac rehabilitation (CR), which is a multicomponent program that educates and supports self-management for the secondary prevention of CHD. CR aims to encourage people to make healthy lifestyle changes for reducing subsequent cardiac events. Lifestyle changes typically include initiating and maintaining regular physical activity, eating a healthy diet, stopping smoking, reducing harmful alcohol intake, and taking medications as per the prescribed regimen [4]. Appropriate self-management is critical for people with CHD to maximize treatment benefits [5].

Despite its benefits, participation in CR has been shown to be inadequate in all countries in which it has been measured [6]. In response, various alternative modes of delivery, including home-based CR and telehealth, have been developed, with similar reductions in cardiovascular risk factors as compared with hospital-based programs [7]. Most recently, mobile health technology has been used to better support the long-term self-management in people with cardiovascular disease [8]. Mobile health has the potential to be an extremely powerful tool for influencing behavior at a population level because it is widely available globally, inexpensive, and allows instant delivery of information [9].

SMS text messaging is the most widely used mobile phone intervention [10,11]. The strongest evidence for SMS text messaging in CHD is from the TEXT ME randomized controlled trial (RCT; n=710) [12], which reported statistically significant positive effects on low-density lipoprotein cholesterol, with more sizable effects on various secondary outcomes. Despite its effectiveness, the study was limited to a single center in Australia and excluded many participants because they did not own a mobile phone, which limited generalizability. TextMe was also evaluated as a standalone strategy; thus, it was unclear whether the intervention was more or less beneficial to those in traditional programs.

To address these limitations, we conducted the Text4HeartII trial across 2 district health boards in Auckland, New Zealand. This trial extended our previous Text4Heart randomized controlled pilot trial (n=123) [13,14], which found a doubling of adherence to lifestyle behaviors at 3 months but not at 6 months. Here, we present findings from a larger effectiveness trial of Text4Heart using an objective measure of medication adherence.

Aims

This study aimed to determine the effectiveness of the Text4HeartII self-management program for improving adherence to medication and lifestyle changes in addition to usual care in people with an acute coronary syndrome (ACS) at 24 and 52 weeks.

Methods

Overview

A two-arm, parallel RCT was conducted in 2 large metropolitan hospitals in Auckland, New Zealand between July 2016 and November 2019. The study was approved by the New Zealand Health and Disability Ethics Committee (15/NTA/205), and the protocol was registered and published before the conclusion of recruitment (Australian New Zealand Clinical Trials Registry, ID: ACTRN12616000422426. Registered, April 1, 2016). The trial was developed and reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) statement (Multimedia Appendix 1). No changes were made to the methods after commencement of the trial.

Participants

Eligible participants were adults with an ACS (including those who had undergone a percutaneous coronary revascularization procedure), clinically stable, able to read English, and able to provide informed consent. Participants were excluded if they had untreated ventricular tachycardia, severe heart failure, life-threatening coexisting disease with life expectancy of less than 1 year, and significant exercise limitations other than cardiovascular disease. Given the high level of mobile phone penetration (there were 6.5 million mobile connections in New Zealand in 2020; population 5.0 million people), mobile phone ownership was not considered in the eligibility criteria [15].

Procedures

Overview

Research nurses recruited participants from 2 metropolitan hospitals in the Auckland region of New Zealand (Auckland City and North Shore Hospitals) before discharge from the hospital following an ACS or post discharge (within 6 weeks) via telephone. Potential participants were contacted in person to determine their interest in the study. Nurses undertook screening to determine their eligibility. Those who met the eligibility criteria were provided with a participant information sheet and consent form. For this trial, informed consent was obtained verbally or in person depending on when the participant

agreed to participate. Interested participants contacted the research nurse to schedule a time for baseline assessments.

The trial was nested within the existing Australian and New Zealand Acute Coronary Syndrome-Quality Improvement (ANZACS-QI) program, which allowed routinely collected data to be accessed for baseline and follow-up assessments. ANZACS-QI [16] is a web-based app deployed nationally to securely gather data on every suspected patient with an ACS in New Zealand and is embedded in more than 90% of hospitals, including those involved in this study. ANZACS-QI provided in-hospital data on all people with an ACS, with risk stratification, diagnosis, investigation management, and complications [17]. Individual behavioral data (diet, exercise, alcohol consumption, and smoking status) were assessed via a telephone interview. In addition, during the 24-week follow-up call, participants were asked to respond to specific questions regarding their overall perceptions of the Text4HeartII program.

Sample Size

A total sample of 330 participants (165 per group) was estimated to provide 80% power at the 5% level of significance (two-sided) to detect an absolute between-group difference of 15% in the proportions of participants adherent to medication at the end of the 24-week intervention (assuming a control rate of 30%). The conservative control rate was based on self-reported medication data from our original Text4Heart study and on previous New Zealand research [18], which found that only 60% of patients had a medication possession ratio (MPR) >0.8 for statins; we believed that this value would be lower if all classes of medication (statins, antihypertensive, and antiplatelet therapy) were included.

Randomization, Allocation Concealment, and Blinding

Upon submission of baseline data, eligible participants were randomly allocated in a 1:1 ratio to the intervention or control groups using block randomization with variable block sizes of 2 or 4, stratified by hospital. The randomization sequence was generated by a biostatistician (YJ). Study investigators (but not participants) were blinded to the intervention allocation during the trial. Primary outcome data (prescribed medication) were obtained via the Ministry of Health National Data Linkage, thereby mitigating any bias associated with self-reported outcome assessments.

Control and Intervention

All participants received usual medical management and were offered center-based CR as per guidelines. CR offered at the participating hospital sites consisted of a 1-hour outpatient education program per week for 6 weeks at a hospital or community center covering a range of topics, such as cardiovascular risk factors, lifestyle changes, and psychosocial support. Patients were also encouraged to attend a 16-session supervised exercise program at the participating hospital or outpatient center. Participants could also participate in usual care CR from the point of discharge to 24 weeks after their cardiac event. In addition to usual care, the intervention group received a 24-week program of automated daily SMS text messages commencing within a week of the baseline assessment. All participants were telephoned at 24 and 52 weeks post randomization to obtain follow-up self-reported outcome data.

Intervention

Text4HeartII comprised a personalized, automated program of self-management that was delivered via SMS text messages over 24 weeks (full details are provided in the study protocol) [19]. The overall goal of the intervention was to have individuals adhere to the New Zealand treatment guidelines for an ACS [20]. Specifically, Text4HeartII included core *Heart Health* content comprising education and support to encourage regular taking of medication, eat a healthy diet (including moderating alcohol consumption), manage stress, and exercise regularly (total 126 messages). Additional SMS text messages were delivered based on the suboptimal behavior participants wanted to modify (eg, physical activity, heart healthy diet, stress management, and stop smoking); each module contained 35 text messages. Modules were discussed with research nurses at baseline to identify participants' preferences. Participants were only able to choose one additional module; however, smokers were prioritized to receive messages providing cessation support. All content was grounded in established psychological (Common Sense Model) [21] and behavior change (social cognitive) theory [22]. The content was focused on modifying people's perceptions of the symptoms, timeline, cause, consequences, personal control over, and the ability of treatment to prevent cardiovascular disease [23] as well as altering the key mediators of behavior change, including self-efficacy, social support, and motivation. The intervention content was based on the original Text4Heart pilot program, with some modifications. In Text4HeartII, we did not include a user website, as this feature was seldom used in the original pilot study [14]. We also revised the message content from weeks 12 to 24 to promote maintenance of the behaviors and relapse prevention.

Participants received a minimum of 1 core heart message per day for 24 weeks, with an additional 35 messages sent over the first 12 weeks; all messages were sent from a centralized server. Messages were sent at times to suit the participants and were personalized with the participant's name. Messages were predominantly unidirectional, but participants were able to text the research team to share their progress if they wanted (eg, goals achieved). Brief training was offered to all participants at enrollment on how to read a text message and how to delete or save messages. No changes were made to the intervention content or delivery during the study period. Examples of the SMS text messages are provided in Textbox 1.

Textbox 1. Examples of SMS text messages.

Heart Health

- T4H: Know your numbers – when is the last time you've had your cholesterol or blood pressure checked? Ask next time you see your GP.
- T4H: High cholesterol or high blood pressure is not good for your heart condition. Your medications will help improve these.
- T4H: Think about your future health. How do you want to feel in 6 months? Try setting small goals with your GP or support person to reach that.

Medication

- T4H: It can be scary to think about the chance of having another heart problem. Taking your pills and changing your lifestyle can lower the risk.
- T4H: It's important to take your medications regularly. To help remember make this part of your daily routine, such as after brushing your teeth.

Physical Activity

- T4H: Changing it up can improve your fitness. Do long slow walks, then head for some hills, then hit the track for speed. Start slow & build.
- T4H: Sometimes, it is hard to exercise when it is raining, try an indoor activity or grab an umbrella and wrap up.

Diet

- T4H: Be wary of low-fat food. Not all are good for you, some are still high in sugar. Read your food labels and compare using the 100g column.
- T4H: To increase your servings of fruit & veg, add a can of tomatoes to your dish or pour frozen veggies into stews/soups/risottos.

Smoking

- T4H: Draw a habit map. When I smoke at work its because..., when I smoke at home its because..., think of the reasons, fix the causes!
- T4H: Find yourself a quit buddy who you can call or txt if you get down. You can make each other feel better with encouragement & share tips.

Stress and Relaxation

- T4H: If you feel discomfort after your heart event, such as feeling angry, sad or withdrawn, consider talking to a health professional.
- T4H: If you feel stressed, close your eyes and imagine a scene where you feel calm. It might be a tropical beach, a forest, or a favorite spot.

Outcomes

All outcomes were assessed at 24 and 52 weeks post randomization. The primary outcome was patient adherence to prescribed medication at 24 weeks, defined as an MPR of 80% or more for 3 medication classes, namely, antiplatelet agent (aspirin), statin, and antihypertensive therapy (angiotensin-converting enzyme inhibitor [ACEI] or angiotensin receptor blocker [ARB] and/or a β-blocker), consistent with the guideline-recommended therapy [20]. To obtain the primary outcome, MPR was calculated for each drug class and then combined to determine adherence across the 3 combined classes. The choice of primary outcome was driven by the positive effect observed on self-reported medication adherence at 6 months in our pilot study and by the ability to provide an objective assessment of medication adherence (namely MPR). The MPR approach has been used successfully in New Zealand to assess statin use [18]; we adapted this approach for the other medication classes. MPR reflects the number of days the drug was assumed to be in a patient's possession (based on dispensed drugs) divided by the number of days spent out of hospital and alive over the assessment period. In a previous study in New Zealand, the possession of medications during 80% or more of follow-up time (ie, MPR ≥0.8) was indicative of maintenance [18]. This proportion was based on post hoc analyses from previous trials on the effect of statin compliance on coronary events and all-cause mortality [24]. To calculate the MPR, community pharmacy dispensing records were linked using the encrypted National Health Index (NHI) via the National Pharmaceuticals Collection database.

Secondary Outcomes

Objective Secondary Outcomes

The MPR for each class of medication (aspirin, statins, ACEI/ARB, and/or β-blockers) was also assessed at 52 weeks using the same approach as used at 24 weeks. Blood pressure, total cholesterol, low-density lipoprotein cholesterol, and high-density lipoprotein cholesterol data were obtained from the ANZACS-QI registry; these were based on routinely collected data.

Self-reported outcomes were measured by a trained research assistant during a telephone call at 24 and 52 weeks. Self-reported medication adherence was assessed using the Morisky 8-item Medication Adherence Scale [25] (0=high, 1-2=medium, and 3-6=low adherence). A license for use was obtained in this study. Adherence to recommended lifestyle behaviors was measured using a composite health behavior score adapted from the European Prospective into Cancer–Norfolk prospective population study [26]. For this study, we used New Zealand–relevant questions to capture all 4 health behaviors. This approach differed slightly from the European Prospective into Cancer–Norfolk study, which used plasma vitamin C >50 mmol/L to indicate fruit and vegetable intake of at least five servings per day. In our study, the

following measures were used to determine participants' health behavior scores:

1. Smoking status was measured using 3 items from a validated smoking history questionnaire [27] and included "have you ever smoked, have you had a puff in the last week and when they quit smoking" (if appropriate).
2. Physical activity level was assessed using items adapted from the New Zealand Health Survey to assess the daily time spent in light-, moderate-, and vigorous-intensity activities [28]. The time spent in moderate-to-vigorous physical activity (MVPA) was calculated and used to indicate adherence to guidelines [29].
3. Alcohol consumption was measured using the alcohol use disorders identification test of alcohol consumption questions [30], a screening tool designed to assess the units of alcohol consumed per week and to identify people who are hazardous drinkers. Index cards referencing standard drink sizes were used to reduce comprehension errors.
4. Fruit and vegetable intake were assessed by 2 New Zealand–specific questions used in the 2006/2007 New Zealand Health Survey (N=12,488, including adults with CHD) [28].

Participants received a score on a 4-point scale for each of the key risk factors, with 1 point each assigned for being a current nonsmoker, meeting physical activity guidelines to achieve some health benefits (≥150 min of MVPA per week), consuming 14 or fewer standard alcoholic drinks per week, and consuming at least five servings of fruits and vegetables per typical day. A total score of 3 or 4 was considered adherence to healthy behaviors. No changes were made to study the outcomes once the trial commenced.

Health-related quality of life was assessed using the European Quality of Life 5 dimensions [31] measure of health status.

Perceptions of Text4HeartII

During the 24-week follow-up telephone call, participants were asked to respond to specific questions about Text4HeartII. Questions were based on those used in the Text4Heart pilot trial and included their perceptions of the program, technical issues experienced, and whether they changed behaviors. Participants who responded that they had changed behaviors were asked to indicate which specific behaviors they had changed.

Adverse Events

All participants were telephoned at 24 and 52 weeks by a trained researcher to determine whether the participants had experienced any serious adverse events during the course of the study. Serious adverse events were reported to a registered medical physician to determine whether they were associated with the study treatment and to determine the course of action (if needed).

Statistical Analysis

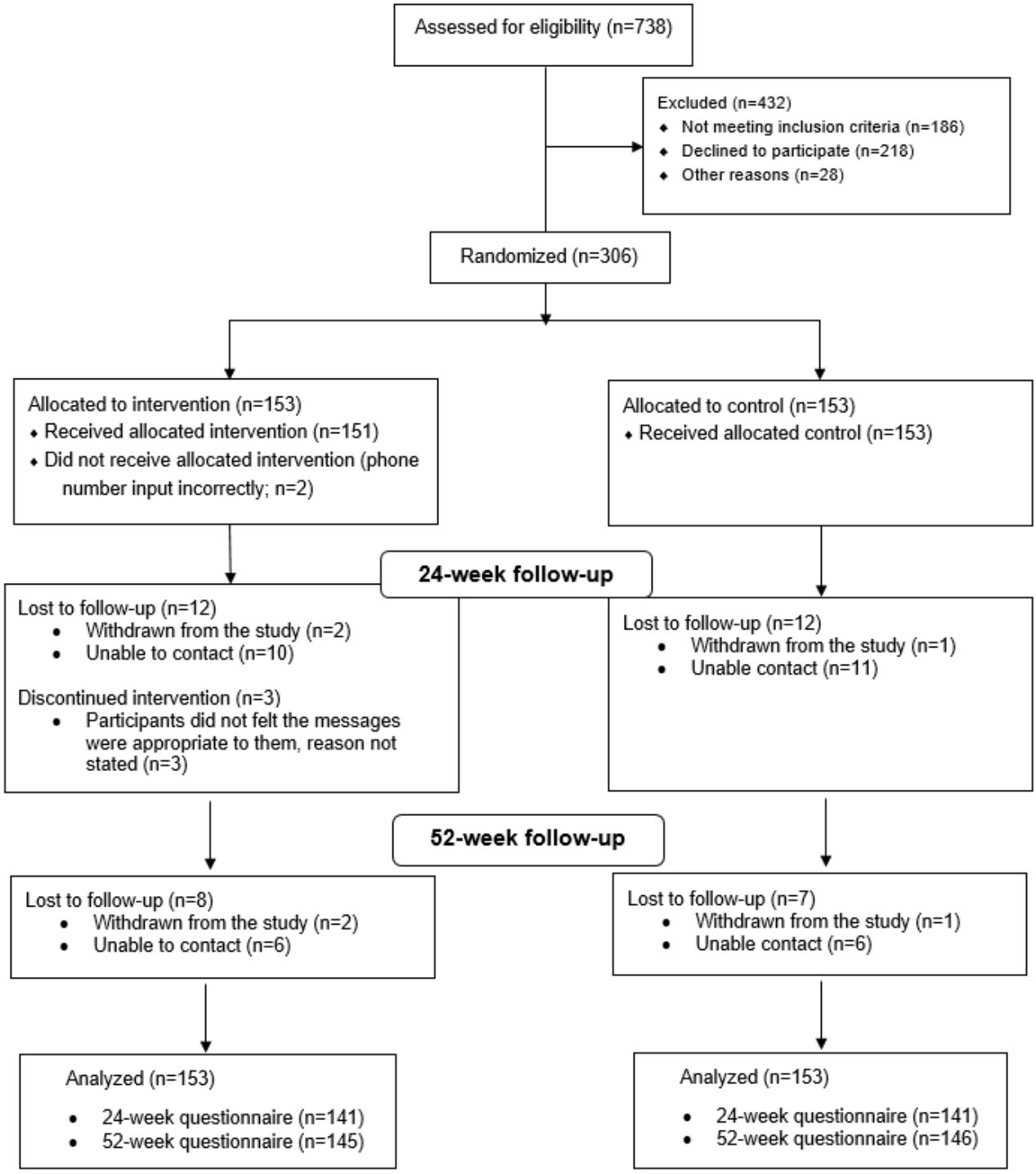
Trial data collected from all eligible participants were linked to the national database using encrypted NHIs for the purpose of analysis. Treatment evaluations were performed according to the intention-to-treat principle. There were no missing data on the MPR (primary outcome), which were obtained via national data linkage. Continuous variables were summarized as mean and SD, and categorical variables were summarized as frequency and percentage. Logistic regression was conducted to evaluate the main treatment effects on medication adherence at 24 and 52 weeks, adjusting for hospital (stratification factor). Odds ratios (ORs) and associated 95% CI were reported at each visit. The same regression models were used for adherence to healthy behaviors at 24 and 52 weeks. The analysis of covariance regression model was used to evaluate the treatment effect on continuous secondary outcomes, adjusting for the baseline value and hospital. Model-adjusted mean differences were reported with 95% CI. Statistical analysis was performed using SAS version 9.4 (SAS Institute Inc.). All statistical tests were two-sided at a 5% significance level.

Results

Overview

Figure 1 presents the flow diagram of the progress through the phases of the trial. A total of 739 people were screened between July 2016 and September 2018. The last trial participant completed the 24-week follow-up on October 2, 2019. Study data collection was completed in November 2019. A total of 306 eligible participants were randomized (153 per group), 282 participants were followed up at 24 weeks (141 per group), and 291 participants were followed up at 52 weeks (intervention: n=145 and control: n=146). All participants had MPR data at both 24 and 52 weeks, which were obtained directly from national medication dispensing records using the linked encrypted NHIs.

Figure 1. Participant flow.



The participants were predominantly male (113/306, 80.3%) and New Zealand European (210/306, 68.6%), with a mean age of 61 years (SD 11) years. The groups were comparable at baseline (Table 1). Most participants were employed full-time (115/306, 37.6%) or retired (132/306, 43.1%). Most participants were married or living with a partner (230/306, 75.2%).

Table 1. Baseline demographic, clinical, and behavioral data of all randomized participants (N=306).

Characteristic	Control (n=153)	Intervention (n=153)
Demographics		
Age (years), mean (SD)	61 (11)	61 (11)
Male, n (%)	113 (73.8)	123 (80.4)
Ethnicity, n (%)		
New Zealand European	102 (66.7)	108 (70.6)
Maori	11 (7.2)	12 (7.8)
Pacific	8 (5.2)	5 (3.3)
Asian	13 (8.5)	11 (7.2)
Other	19 (12.4)	17 (11.1)
Employment, n (%)		
Self-employed	33 (21.6)	29 (18.9)
Full-time employment	55 (35.9)	60 (39.2)
Part-time employment	9 (5.9)	5 (3.3)
Retired	46 (30.1)	46 (30.1)
Other (full-time homemaker, student, unemployed, or beneficiary)	10 (6.5)	13 (8.5)
Marital status, n (%)		
Married or living with partner	124 (81)	106 (69.3)
Separated, divorced, or widowed	24 (15.7)	33 (21.6)
Never married	5 (3.3)	14 (9.1)
Clinical data, mean (SD)		
BMI (kg/m^2)	29.21 (5.6)	29.74 (5.9)
Cholesterol (mmol/L)		
Total	4.74 (1.2)	5.07 (1.3)
HDL-C[a]	1.19 (0.3)	1.20 (0.5)
LDL-C[b]	2.63 (1.1)	2.99 (1.2)
Systolic blood pressure (mm Hg)	129 (16)	129 (18)
Diastolic blood pressure (mm Hg)	74 (10)	75 (13)
Behavioral risk factors		
Meeting MVPA[c] guidelines, n (%)	70 (46)	75 (49)
Not smoking, n (%)	140 (91)	138 (90)
Number of drinks ≤14 per week, n (%)	132 (86)	133 (87)
Fruit and vegetables serves ≥5 week, n (%)	84 (55)	71 (46)
HRQoL[d]—health state score, mean (SD)	64 (22)	62 (22)

[a]HDL-C: high-density lipoprotein cholesterol.

[b]LDL-C: low-density lipoprotein cholesterol.

[c]MVPA: moderate-to-vigorous physical activity.

[d]HRQoL: health-related quality of life.

Adherence to the 3 medication classes (aspirin, statin, and any antihypertensive) in the intervention group was worse than that in the control group at 24 weeks (OR 0.60, 95% CI 0.38-0.96; *P*=.03). Differences in adherence to all 4 medication classes (aspirin, statin, β-blocker, and ACEI/ARB; OR 0.67, 95% CI 0.42-1.05; *P*=.08) and for individual classes of aspirin, statins, and β-blockers (Table 2) were not statistically significant, except for adherence to ACEI/ARB (OR 0.49, 95% CI 0.30-0.82; *P*=.006), which was worse in the intervention group than in the control group.

Table 2. The effect of intervention on medication adherence at 24- and 52-week follow-up (full cohort).

Adherence to	24 weeks (N=306)				52 weeks (N=306)			
	Control (n=153), n (%)	Intervention (n=153), n (%)	Adjusted OR[a] (95% CI)	*P* value	Control (n=153), n (%)	Intervention (n=153), n (%)	Adjusted OR (95% CI)	*P* value
All 3 drug classes (aspirin, statin, and any BP[b]-lowering drug): primary outcome	105 (68.6)	87 (56.8)	0.60 (0.38-0.96)	.03	104 (67.9)	83 (54.2)	0.56 (0.35-0.89)	.01
All 4 drug classes (aspirin, statin, β-blocker, and ACEI[c]/ARB[d])	71 (46.4)	56 (36.6)	0.67 (0.42-1.05)	.08	70 (45.7)	56 (36.6)	0.68 (0.43-1.08)	.11
Statin	131 (85.6)	122 (79.7)	0.66 (0.36-1.20)	.18	129 (84.3)	119 (77.7)	0.65 (0.36-1.16)	.15
Aspirin	124 (81.0)	122 (79.7)	0.92 (0.52-1.62)	.78	123 (80.3)	119 (77.7)	0.85 (0.49-1.49)	.58
β-Blocker	103 (67.3)	100 (65.3)	0.92 (0-1.48)	.73	102 (66.6)	89 (58.1)	0.69 (0.43-1.11)	.13
ACEI/ARB	119 (77.7)	97 (63.4)	0.49 (0.30-0.82)	.006	123 (80.3)	97 (63.4)	0.42 (0.25-0.71)	.001
BP-lowering drugs (ACEI/ARB and/or β-blocker)	137 (89.5)	122 (79.7)	0.46 (0.24-0.88)	.02	139 (90.8)	113 (73.8)	0.28 (0.15-0.55)	<.001

[a]OR: odds ratio; odds ratio compares the estimated odds between intervention and control groups.

[b]BP: blood pressure.

[c]ACEI: angiotensin-converting enzyme inhibitor.

[d]ARB: angiotensin II receptor blockers.

Similar results were found for medication adherence at 52 weeks (secondary outcome), with higher adherence to 3 medication classes in the control group than in the intervention group (OR 0.56, 95% CI 0.35-0.89; *P*=.01). The difference in adherence to the 4 medication classes was not significant (OR 0.68, 95% CI 0.43-1.08; *P*=.10). Similarly, there were no significant differences for individual classes of aspirin, statins, and β-blockers (Table 2), but adherence to ACEI/ARB was less in the intervention group than in the control group (OR 0.42, 95% CI 0.25-0.71; *P*=.001). These data are mirrored to some extent by the self-reported medication adherence data. At 24 weeks, there was a small effect on self-reported medication adherence, which favored the control group (adjusted mean difference 0.15; 95% CI −0.15 to 0.45; *P*=.33), with a larger effect observed at 52 weeks (adjusted mean difference 0.30; 95% CI 0.01-0.59; *P*=.04.

The composite measure of adherence with lifestyle behaviors was similar for both the intervention and control groups at 24 weeks (OR 1.11, 95% CI 0.65-1.90; *P*=.70) and 52 weeks (OR 0.97, 95% CI 0.58-1.62; *P*=.90; Table 3). In terms of individual behaviors, the number of participants who did not smoke at 24 weeks was higher in the intervention group (n=138) than in the control group (n=132; OR 5.75, 95% CI 1.08-30.61; *P*=.04). No other differences were observed between these groups. At 52 weeks, there was a trend for differences in time spent in vigorous-intensity activity (mean difference 44.8 min; 95% CI −3.9 to 93.5; *P*=.07) and MVPA (mean difference 100.3 min; 95% CI −16.6 to 217.3; *P*=.09), which favored the intervention group. A total of 75 serious adverse events were reported during the trial; however, none of them were related to the study. Only 1 participant died during the study period.

Table 3. The effect of intervention on adherence to lifestyle risk factors (secondary outcomes) at 24- and 52-week follow-up.

Adherence to	24 weeks (n=282)						52 weeks (n=291)					
	Control		Intervention		Adjusted OR[a] (95% CI)	*P* value	Control		Intervention		Adjusted OR (95% CI)	*P* value
	Sample size, n	Participant, n (%)	Sample size, n	Participant, n (%)			Sample size, n	Participant, n (%)	Sample size, n	Participant, n (%)		
Lifestyle behaviors (composite)	141	94 (66.7)	141	100 (70.9)	1.11 (0.65-1.90)	.70	146	95 (65.1)	145	96 (66.2)	0.97 (0.58-1.62)	.90
Physical activity guideline	141	68 (48.3)	140	79 (56.4)	1.43 (0.87-2.35)	.16	146	80 (54.8)	145	88 (60.7)	1.25 (0.75-2.09)	.39
Smoking cessation	141	132 (93.6)	141	138 (97.8)	5.75 (1.08 - 30.61)	.04	146	139 (95.2)	145	139 (95.7)	1.38 (0.39-4.88)	.62
Low alcohol consumption	140	132 (94.3)	141	130 (92.2)	0.68 (0.24-1.87)	.45	145	141 (97.2)	145	136 (93.8)	0.33 (0.09-1.28)	.11
Fruit and vegetable guidelines	141	67 (47.5)	141	64 (45.4)	1.01 (0.61-1.66)	.98	144	64 (44.4)	145	54 (37.2)	0.81 (0.49-1.33)	.40

[a]OR: odds ratio; odds ratio compares the estimated odds between intervention and control groups.

Perceptions of Text4HeartII

Table 4 shows participants' responses to questions regarding the Text4HeartII intervention. Overall, most participants (136/139, 97.8%) did not experience any technical issues and would recommend the program to others. Most (114/138, 82.6%) participants felt that the program helped them manage their heart disease, but perceptions of whether the program influenced the behavior were mixed.

Table 4. Participants' responses on Text4HeartII (n=139).

Question	Sample size, n	Participants, n (%)
"Did you have any technical problems with the program?"	139	136 (97.8)
"Would you recommend the program to other people who have had a heart event?"	138	134 (97.1)
"Did taking part in this program help you learn about heart condition?"	138	86 (62.3)
"Did taking part in this program help you in the recovery from your heart condition?"	138	114 (82.6)
"**Did taking part in this program help you change your behaviors?"**	138	78 (56.5)
"**If yes, which behaviors did you change?"**	78	
"I became physically active"		41 (52.5)
"I ate more fruit and vegetables"		34 (43.6)
"I ate less saturated fat"		21 (26.9)
"I took my medication regularly"		19 (24.3)
"I drank less alcohol"		10 (12.8)
"I ate less salt"		10 (12.8)
"I lowered my level of stress"		9 (11.5)
"I lost weight"		7 (9.0)
"I stopped smoking"		5 (6.4)
"I ate more healthy fat"		4 (5.1)
"I had regular GP[a] checks"		3 (3.8)
"I watched less TV[b]"		0 (0)
"I got more adequate sleep"		0 (0)

[a]GP: general practitioner.

[b]TV: television.

Discussion

Principal Findings

The Text4HeartII trial extended our pilot trial to determine the effectiveness of an SMS text messaging–based intervention to improve adherence to medication and lifestyle behaviors at both 24 and 52 weeks. Overall, we found no evidence to support the effectiveness of the program on dispensed medication or on adherence to a composite measure of lifestyle change over and above usual care.

The strengths of this trial were the RCT design and the objective assessment of medication adherence. Our study addressed criticisms of previous SMS text messaging trials [32,33], including enhanced generalizability. Our trial was of sufficient duration to elicit behavior change, and the follow-up assessments (24 and 52 weeks) were long enough to determine the sustained effect of the intervention. An additional strength of this trial was the complete data on the primary outcome. A possible limitation could be using MPR as an outcome, which reflects the number of days a drug was assumed to be in a person's possession (based on dispensed drugs) rather than a measure of whether a person took their medication or not. Nevertheless, previous studies have used MPR as a proxy measure of medication adherence [20].

Comparison With Other Studies

The lack of a positive effect on mediation adherence was surprising and contrary to previous research. A meta-analysis of 16 RCTs (N=2742) showed that SMS text messaging significantly improved medication adherence across a range of conditions (OR 2.11, 95% CI 1.52-2.93; $P<.001$) [34]. The effect was not sensitive to study characteristics (intervention duration or type of disease) or text message characteristics (personalization, 2-way communication, or daily text message frequency). The most commonly used method to assess adherence was self-report (9 studies), followed by the medication event monitoring system (4 studies).

Our findings differ from previous SMS text messaging trials in people with cardiovascular disease. A recent review (Text2PreventCVD) involving 9 trials (N=3779) and individual participant data (5 trials; n=2612) meta-analysis of SMS text messaging interventions demonstrated positive effects across a host of risk factors (BMI, systolic blood pressure, and diastolic blood pressure) [35]. Without directly comparing the messages within the respective interventions, it is difficult to know where the differences in effect may exist. However, a recent systematic description and comparison of the development processes of the interventions included in that systematic review highlighted that the Text4HeartII development process was comparable with that of previous trials [36]. Moreover, our Text4Heart pilot trial was included in the Text2PreventCVD systematic review

and comparison of development processes. Like similar trials, the development of Text4HeartII involved consultation with experts, users, or other stakeholders; was based on literature reviews; included relevant theory; and involved primary research with end users [36]. The behavior change techniques used in this study were similar to those used in previous studies [36].

Given the existing body of evidence supporting SMS text messaging on medication adherence, it is clear that Text4HeartII did not have the desired effect on the population and setting in this trial. This may be attributed to personal factors associated with our cohort; randomized participants appeared to be more motivated to adhere than those in previous studies. The results showed that the difference between the control and intervention groups was relatively small and driven predominantly by MPR for antihypertensive medication. Despite similar days in hospital between groups at 24 weeks (approximately 1 day), adherence to all medication classes for both groups was higher than our proposed control rate (30%); approximately 41.5% (127/306) of the participants had an MPR >0.8. Our original control rate was based on previous research, which found that 59.3% (8028/13,520) of patients had an MPR >0.8 for statins only [18]. However, in this study, adherence to statins was higher; 82.6% (253/306) of participants had an MPR >0.8. Second, in our study, participants were recruited by cardiac nurses, and all participants were offered CR services, with good access to follow-up. The lack of marked differences between groups may have reflected organizational factors with higher levels of available support (including CR) at the 2 participating metropolitan hospitals. Thus, the effectiveness of SMS text messaging interventions may be strongly influenced by the population and context in which they are applied. This highlights the need to better understand the context in which SMS text messaging interventions are delivered and how individual- and organizational-level factors may affect the adoption, implementation, and effectiveness of these types of interventions. It also emphasizes the importance of evaluating interventions in the setting they are likely to be used before widespread adoption. Third, it is also possible that our text messages were not potent enough to evoke changes in medication adherence over and above usual cardiac care. Many previous SMS text messaging interventions have targeted single behaviors, such as smoking behavior, exercise, and medication adherence [32,34,37,38]. Thus, the lack of potency may have resulted from our approach to target multiple behaviors rather than targeting medication adherence alone. This is consistent with a previous qualitative research, which showed that conversations about changing multiple health behaviors were perceived to be overwhelming for patients and difficult to implement for health care professionals [39]. The lack of potency hypothesis is in line with qualitative responses from participants in this study, which showed that although 82.6% (114/138) of intervention participants responded that participating in this program had helped them recover from their heart condition, only 24.3% (19/78) felt the messages helped them to take their medications regularly. Furthermore, 56.5% (78/138) of the intervention participants responded that the program helped them change their health-related behaviors. Fourth, it is possible that the SMS text messaging program did not fully address the needs of participants [40]. Previous authors have highlighted the need for more personalized messages, depending on diagnosis (atherosclerosis, spontaneous coronary artery dissection, and Takotsubo cardiomyopathy) and key risk factors such as current smoking status. Although the proposed message library developed by Marshall et al [40] was similar to ours in terms of advice on a graduated exercise program, nutrition, smoking cessation, stress management, and medication, their proposed content also included messages around support groups and information about respective diagnoses. Patients could also send a message, which prompted a response from a CR nurse. The results of that study have not been published and were therefore not available for direct comparison with this study.

Surprisingly, the effects observed on lifestyle behaviors differed from those in our original pilot study [14]. In this study, we showed no clear differences in our composite measure of lifestyle change or individual behaviors, beyond those reported. The lack of effect may be attributed to the following issues. First, it is possible that there was a ceiling effect, with approximately 59.1% (181/306) of participants reporting adherence to 3 or more lifestyle behaviors at baseline. This could reflect self-reported bias, which was evident in the measurement of physical activity. Second, our measure of physical activity in this study differed from that in the pilot study. In this study, we used 3 items to capture the time spent in light-, moderate-, and vigorous-intensity activities. This approach was used to provide a better metric for adherence to guidelines for physical activity in adults (≥150 min of MVPA per day). Descriptive data suggested substantial overreporting of physical activity levels at all time points. For example, participants at baseline in both groups reported more than 300 minutes of daily MVPA, which is considerably higher than the standard population estimate for adults in New Zealand. Third, the lack of effect may have been related to the fact that both groups had well-managed risk factors at baseline (eg, few smokers, well-managed blood pressure), which limited the potential for change. Fourth, the lack of effect may also be attributed to changes made from the original intervention, which involved removing the website, which allowed people to set goals, review previous text messages, and access other resources related to CHD. We decided to remove the website component as it was infrequently used, and fewer than half of the participants in the pilot study felt using a website was a good way to deliver the program [14]. In the original study, we also issued pedometers and allowed people to track step counts, which was not a feature of this study. These features may have allowed for more interaction with the program but were removed because of pragmatic reasons (costs and distribution) related to the potential to scale as a national program.

Future Research

Despite the lack of effect observed in our trial, SMS text messaging as an intervention has the potential to improve outcomes in people with CHD and other conditions. Findings from this study suggest that the context in which SMS text messaging interventions are delivered is important to consider and may have a significant impact on whether an intervention is effective or not. Future studies need to explore both individual- and organizational-level factors that may affect the

adoption, implementation, and effectiveness of such interventions. For example, the RE-AIM (Reach, Effectiveness, Adoption, Implementation and Maintenance) framework [41,42] is a widely used evaluative framework for guiding the evaluation and reporting of health intervention effectiveness. It emphasizes collecting information about the reach, effectiveness, adoption, implementation, and maintenance of an intervention across both individual- and setting- or staff-level variables. RE-AIM aligns with systems-based approaches and allows for the assessment of vertical (eg, adoption decisions within a given organization) and horizontal (eg, adoption across different sectors) components. The application of RE-AIM or a similar framework would help in providing a clear understanding of the key factors that may affect the effectiveness of future text-based messaging interventions.

Conclusions

There was no evidence to support the effectiveness of Text4HeartII on dispensed medication or adherence to a favorable lifestyle over and above usual care. In its current form, Text4HeartII cannot be used to augment existing services. The findings of this study are in contrast with those of previous studies and highlight the importance of evaluating interventions in the setting they are likely to be used before widespread adoption. Changes to the intervention program are warranted to justify its future implementation.

Acknowledgments

The authors would like to thank all the participants involved in the Text4HeartII trial. The authors would also like to acknowledge the participating hospitals (Auckland, North Shore, and Waitakere) and the cardiac staff who assisted with recruitment and conduct of the trial. Finally, the authors wish to thank the Vascular Informatics Using Epidemiology and the Web and the ANZACS-QI teams for providing access to data for this study. The Text4HeartII trial was funded by the Health Research Council of New Zealand and the National Heart Foundation.

Authors' Contributions

RM conceived the idea, procured funding, was responsible for the overall conduct of the trial, and drafted the manuscript. YJ performed data analyses. LD was involved in the original Text4Heart pilot trial on which this study was based. All investigators were involved in the conduct of the trial, contributed to the interpretation of the findings, and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) statement checklist.
[PDF File (Adobe PDF File), 182 KB - mhealth_v9i6e24952_app1.pdf]

References

1. Cardiovascular diseases (CVDs). World Health Organization. 2011. URL: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2021-05-21]
2. Australian Institute of Health and Welfare. Cardiovascular disease, diabetes and chronic kidney disease: Australian facts: prevalence and incidence. Canberra, Australia: Australian Institute of Health and Welfare; 2014:1-100.
3. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics - 2015 update: a report from the American Heart Association. Circulation 2015 Jan 27;131(4):29-322. [doi: 10.1161/CIR.0000000000000152] [Medline: 25520374]
4. Taylor R, Kirby B. The evidence base for the cost effectiveness of cardiac rehabilitation. Heart 1997 Jul 01;78(1):5-6 [FREE Full text] [doi: 10.1136/hrt.78.1.5] [Medline: 9290390]
5. Riegel B, Moser DK, Buck HG, Dickson VV, Dunbar SB, Lee CS, et al. Self‐care for the prevention and management of cardiovascular disease and stroke. J Am Heart Assoc 2017 Sep 22;6(9):e006997. [doi: 10.1161/jaha.117.006997]
6. Bethell HJ, Evans JA, Turner SC, Lewin RJ. The rise and fall of cardiac rehabilitation in the United Kingdom since 1998. J Public Health (Oxf) 2007 Mar;29(1):57-61. [doi: 10.1093/pubmed/fdl091] [Medline: 17189295]
7. Clark RA, Conway A, Poulsen V, Keech W, Tirimacco R, Tideman P. Alternative models of cardiac rehabilitation: a systematic review. Eur J Prev Cardiol 2015 Jan 13;22(1):35-74. [doi: 10.1177/2047487313501093] [Medline: 23943649]
8. Pfaeffli Dale L, Dobson R, Whittaker R, Maddison R. The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: a systematic review. Eur J Prev Cardiol 2016 May 21;23(8):801-817. [doi: 10.1177/2047487315613462] [Medline: 26490093]
9. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010 Mar 30;32(1):56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
10. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]

11. 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] [Medline: 19135907]
12. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. J Am Med Assoc 2015 Sep 22;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
13. Dale LP, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Improving coronary heart disease self-management using mobile technologies (Text4Heart): a randomised controlled trial protocol. Trials 2014 Mar 04;15(1):71 [FREE Full text] [doi: 10.1186/1745-6215-15-71] [Medline: 24588893]
14. Dale LP, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Text message and internet support for coronary heart disease self-management: results from the Text4Heart randomized controlled trial. J Med Internet Res 2015 Oct 21;17(10):e237 [FREE Full text] [doi: 10.2196/jmir.4944] [Medline: 26490012]
15. Kemp S. Digital 2020: New Zealand. 2020. URL: https://datareportal.com/reports/digital-2020-new-zealand [accessed 2020-11-30]
16. Kerr A, Williams MJ, White H, Doughty R, Nunn C, Devlin G, et al. The All New Zealand Acute Coronary Syndrome Quality Improvement Programme: implementation, methodology and Cohorts (ANZACS-QI 9). N Z Med J 2016 Aug 05;129(1439):23-36. [Medline: 27507719]
17. Kerr AJ, McLachlan A, Furness S, Broad J, Riddell T, Jackson R, et al. The burden of modifiable cardiovascular risk factors in the coronary care unit by age, ethnicity, and socioeconomic status - PREDICT CVD-9. N Z Med J 2008 Nov 07;121(1285):20-33. [Medline: 19079434]
18. Thornley S, Marshall R, Chan WC, Kerr A, Harrison J, Jackson G, et al. Four out of ten patients are not taking statins regularly during the 12 months after an acute coronary event. Eur J Prev Cardiol 2012 Jun 14;19(3):349-357. [doi: 10.1177/1741826711403069] [Medline: 21450568]
19. Maddison R, Stewart R, Doughty R, Scott T, Kerr A, Benatar J, et al. Text4Heart II - improving medication adherence in people with heart disease: a study protocol for a randomized controlled trial. Trials 2018 Jan 25;19(1):70 [FREE Full text] [doi: 10.1186/s13063-018-2468-z] [Medline: 29370829]
20. New Zealand Guidelines Group. Assessment and Management of Cardiovascular Risk. Wellington: New Zealand Guidelines Group; 2003.
21. Leventhal H, Meyer D, Nerenz R. The common sense representation of illness danger. In: Contributions to Medical Pyschology. United Kingdom : Pergamon Press; 1980:7-30.
22. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [doi: 10.1037/0003-066x.44.9.1175] [Medline: 2782727]
23. Broadbent E, Ellis CJ, Thomas J, Gamble G, Petrie KJ. Further development of an illness perception intervention for myocardial infarction patients: a randomized controlled trial. J Psychosom Res 2009 Jul;67(1):17-23. [doi: 10.1016/j.jpsychores.2008.12.001] [Medline: 19539814]
24. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A clinical trial of the effects of dietary patterns on blood pressure. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/nejm199704173361601]
25. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10(5):348-354 [FREE Full text] [doi: 10.1111/j.1751-7176.2008.07572.x] [Medline: 18453793]
26. Khaw K, Wareham N, Bingham S, Welch A, Luben R, Day N. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. PLoS Med 2008 Jan 08;5(1):e12 [FREE Full text] [doi: 10.1371/journal.pmed.0050012] [Medline: 18184033]
27. Maddison R, Roberts V, Bullen C, McRobbie H, Jiang Y, Prapavessis H, et al. Design and conduct of a pragmatic randomized controlled trial to enhance smoking-cessation outcomes with exercise: the Fit2Quit study. Ment Health Phys Act 2010 Dec;3(2):92-101. [doi: 10.1016/j.mhpa.2010.09.003]
28. Methodology report for the 2006/07 New Zealand Health Survey. Ministry of Health. 2008. URL: https://www.moh.govt.nz/NoteBook/nbbooks.nsf/0/38A0B04E793F5F1ECC257DC60008B736/$file/methodology-report-for-nz-health-survey-08.pdf [accessed 2021-05-24]
29. How much activity is recommended? New Zealand Ministry of Health. 2019. URL: https://www.health.govt.nz/your-health/healthy-living/food-activity-and-sleep/physical-activity/how-much-activity-recommended [accessed 2020-12-09]
30. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
31. Rabin R, de Charro F. EQ-5D. APA PsycTests 2001;33(5):1-43. [doi: 10.1037/t57920-000]
32. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul;378(9785):49-55. [doi: 10.1016/s0140-6736(11)60701-0]

33. Kaplan RM, Stone AA. Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Annu Rev Psychol 2013 Jan 03;64(1):471-498. [doi: 10.1146/annurev-psych-113011-143736] [Medline: 22994919]
34. 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 01;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
35. Islam SM, Chow C. A13510 Mobile phone text-messaging interventions aimed to prevent cardiovascular diseases (Text2PreventCVD). J Hypertens 2018;36:314-315. [doi: 10.1097/01.hjh.0000549284.06955.8d]
36. Ricci-Cabello I, Bobrow K, Islam SM, Chow CK, Maddison R, Whittaker R, et al. Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191 [FREE Full text] [doi: 10.2196/12191] [Medline: 30924790]
37. Maddison R, Pfaeffli L, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
38. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug 01;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
39. Alageel S, Gulliford MC, McDermott L, Wright AJ. Implementing multiple health behaviour change interventions for cardiovascular risk reduction in primary care: a qualitative study. BMC Fam Pract 2018 Oct 30;19(1):171 [FREE Full text] [doi: 10.1186/s12875-018-0860-0] [Medline: 30376826]
40. Marshall W, Galaz E, Belz L, Benatar J. A text/email messaging service for cardiac rehabilitation. Heart Lung Circ 2018;27:38. [doi: 10.1016/j.hlc.2018.05.186]
41. Glasgow RE, Klesges LM, Dzewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. Health Educ Res 2006 Oct 15;21(5):688-694. [doi: 10.1093/her/cyl081] [Medline: 16945984]
42. Glasgow RE, McKay H, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Educ Couns 2001 Aug;44(2):119-127. [doi: 10.1016/s0738-3991(00)00186-5]

Abbreviations

ACEI: angiotensin-converting enzyme inhibitor
ACS: acute coronary syndrome
ANZACS-QI: Australian and New Zealand Acute Coronary Syndrome-Quality Improvement
ARB: angiotensin receptor blocker
CHD: coronary heart disease
CONSORT: Consolidated Standards of Reporting Trials
CR: cardiac rehabilitation
MPR: medication possession ratio
MVPA: moderate-to-vigorous physical activity
NHI: National Health Index
OR: odds ratio
RCT: randomized controlled trial
RE-AIM: Reach, Effectiveness, Adoption, Implementation and Maintenance

Edited by R Kukafka; submitted 11.10.20; peer-reviewed by S Lear, J Redfern; comments to author 30.10.20; revised version received 09.12.20; accepted 03.05.21; published 09.06.21.

Please cite as:
Maddison R, Jiang Y, Stewart R, Scott T, Kerr A, Whittaker R, Benatar J, Rolleston A, Estabrooks P, Dale L
An Intervention to Improve Medication Adherence in People With Heart Disease (Text4HeartII): Randomized Controlled Trial
JMIR Mhealth Uhealth 2021;9(6):e24952
URL: https://mhealth.jmir.org/2021/6/e24952
doi:10.2196/24952
PMID:34106081

Review

Efficacy of Short Message Service Text Messaging Interventions for Postoperative Pain Management: Systematic Review

Christoph Buck[1,2*], Dipl-Kfm, MSc, PhD; Christian Keweloh[3*], BA, MSc; Adam Bouras[4*], BSc, MS, PhD; Eduardo J Simoes[4*], MD, MSc, MPH

[1]School of Business and Economics, Philipps-Universität Marburg, Marburg, Germany

[2]Project Group Business and Information Systems Engineering, Fraunhofer Institute for Applied Information Technology FIT, Bayreuth, Germany

[3]UNITY AG, Büren, Germany

[4]School of Medicine, University of Missouri, Columbia, MO, United States

[*]all authors contributed equally

Corresponding Author:
Christoph Buck, Dipl-Kfm, MSc, PhD
School of Business and Economics
Philipps-Universität Marburg
Am Plan 1
Marburg, 35037
Germany
Phone: 49 064212823914
Email: christoph.buck@wiwi.uni-marburg.de

Abstract

Background: Addiction to opiates and synthetic opioids poses a major threat to public health worldwide, with pharmaceutical opioids prescribed to manage pain constituting the main problem. To counteract this threat, suitable pain management strategies should be implemented in health care. Monitoring pain management seems to be feasible using telemedicine with a certain degree of resource intensity and digitization. As a communication channel for this type of monitoring, SMS appears to be a valid alternative.

Objective: The aim of this systematic literature review was to (1) provide information on the state of research regarding postoperative pain management via SMS, (2) establish a basic understanding of SMS-based pain management, and (3) provide insight into the feasibility of these management strategies. The research question was as follows: Is postoperative pain management feasible and effective utilizing SMS?

Methods: A systematic literature review was performed mainly following the PRISMA guidelines and another guide on performing a systematic literature review for information systems–related research. A search string was developed based on the objectives and research question, and eight databases were searched.

Results: The initial search resulted in 2083 records, which could be narrowed down by applying various exclusion criteria. Thereby, 11 articles were identified as relevant, which were accordingly analyzed and evaluated by full-text screening. In all articles, pain management interventions were performed using SMS communication between health care professionals and patients or their legal guardians. A prospective approach was predominantly chosen as the study design (91%) with the leading research objective of determining the intervention's feasibility (73%). The primary reason for sending SMS messages was to monitor patients (64%). Overall, the use of SMS improved adherence, acceptance, and satisfaction regarding postoperative pain management. With an average response rate of approximately 89.5% (SD 3.8%), the reliability of SMS as a communication and monitoring tool was further emphasized. This response rate is significantly higher than that for email interventions (66.63%, $P<.001$).

Conclusions: This study provides a comprehensive picture of the current status on postoperative pain management by SMS. Communication via SMS was beneficial in all interventions, even preoperative. Six SMS interventions could be certified by the respective institutional review board and three were Health Insurance Portability and Accountability Act–compliant. Therefore, the results of this study could be leveraged to address the opioid epidemic. Overall, the research question could be confirmed. Future research should extend this systematic literature review regarding preoperative pain management. Based on these findings, a pre- and postoperative communication model should be developed to address the opioid epidemic effectively.

(JMIR Mhealth Uhealth 2021;9(6):e20199) doi:10.2196/20199

KEYWORDS

systematic literature review; pain management; opioid; short message service (SMS); postoperative

Introduction

Background

Globally, the increasing use of opiates and synthetic opioids poses a major threat to public health [1,2]. In 2017, approximately 53 million people took opioids at least once in the past year, with the highest prevalence of nonmedical opioid use estimated in North America [2]. More than 700,000 people died from drug overdoses in the United States between 1999 and 2017, and approximately two-thirds of these cases involved opioids [3,4]. The leading causes of this epidemic are opioid misuse, an overall increase in opioid prescriptions, shifted patient expectations, inadequate medical education and practice, insufficient guidelines, and the highly addictive nature of opioids [1,3-8]. The major opioids of concern remain pharmaceutical opioids used for pain control [1,2], which are typically prescribed for postoperative pain management [9]. Therefore, suitable pain management strategies need to be developed and implemented to adequately address and counteract this opioid epidemic [10-12].

Almost every aspect of these pain management strategies, and the efficiency and quality of health care rely on effective communication [13,14]. Poor health care professional–patient interactions can lead to adverse clinical outcomes, insufficient patient understanding, poor patient compliance, and consequently negative outcomes [15,16]. Improving communication and implementing postoperative monitoring appears to be a practical approach, since less than half of patients report adequate postoperative pain relief [17]. However, resource-intensive pain management is difficult to implement in the health care sector due to the constant and increasing pressure to provide patient care most efficiently and as cost-effectively as possible [18]. These demands could be fulfilled by increasing telemedicine interventions and measures [19-21].

Mobile health, delivered through mobile instant messaging apps or SMS texting, has particular potential in this regard. SMS is utilized more frequently [22], as SMS communication provides various features and proven benefits for health care applications. Unlike mobile instant messaging apps, no smartphone or internet access is required for SMS [22-24]. In the United States, 96% of the population already own a mobile phone and 97% of smartphone users have sent at least one SMS message within the week. With approximately 6 billion SMS messages sent daily, it is the most popular and widely used communication feature [25,26]. Furthermore, SMS is a low-cost, provider-independent, scalable, ubiquitous, reliable, secure, widely accepted, and simple communication means [22,23,26,27].

Objective and Structure of the Study

To address the ongoing opioid epidemic, pain management combined with SMS as a communication medium appears to be very viable, whereby the postoperative phase seems to be unusually decisive. Therefore, the aim of this study was to structure the current state of the literature regarding postoperative pain management via SMS. To our knowledge, there is currently no specific literature review on SMS-based pain management and no synthesized results. Accordingly, this study examined the following research question: Is postoperative pain management feasible and effective utilizing SMS? By answering the research question through a systematic literature review, a conceptual framework for future research is provided.

To gain a valid answer to the research question, the paper is structured as follows. In the Methods, we describe the process of performing the systematic literature review, along with a detailed description of the specific selection and exclusion criteria. The results of the selected literature are compiled accordingly in the Results. The Discussion explains the principal insights from the included studies, along with the limitations of this review. Finally, we provide recommendations for action based on the conclusions, and highlight research gaps for researchers, clinicians, and other health care professionals.

Methods

Design

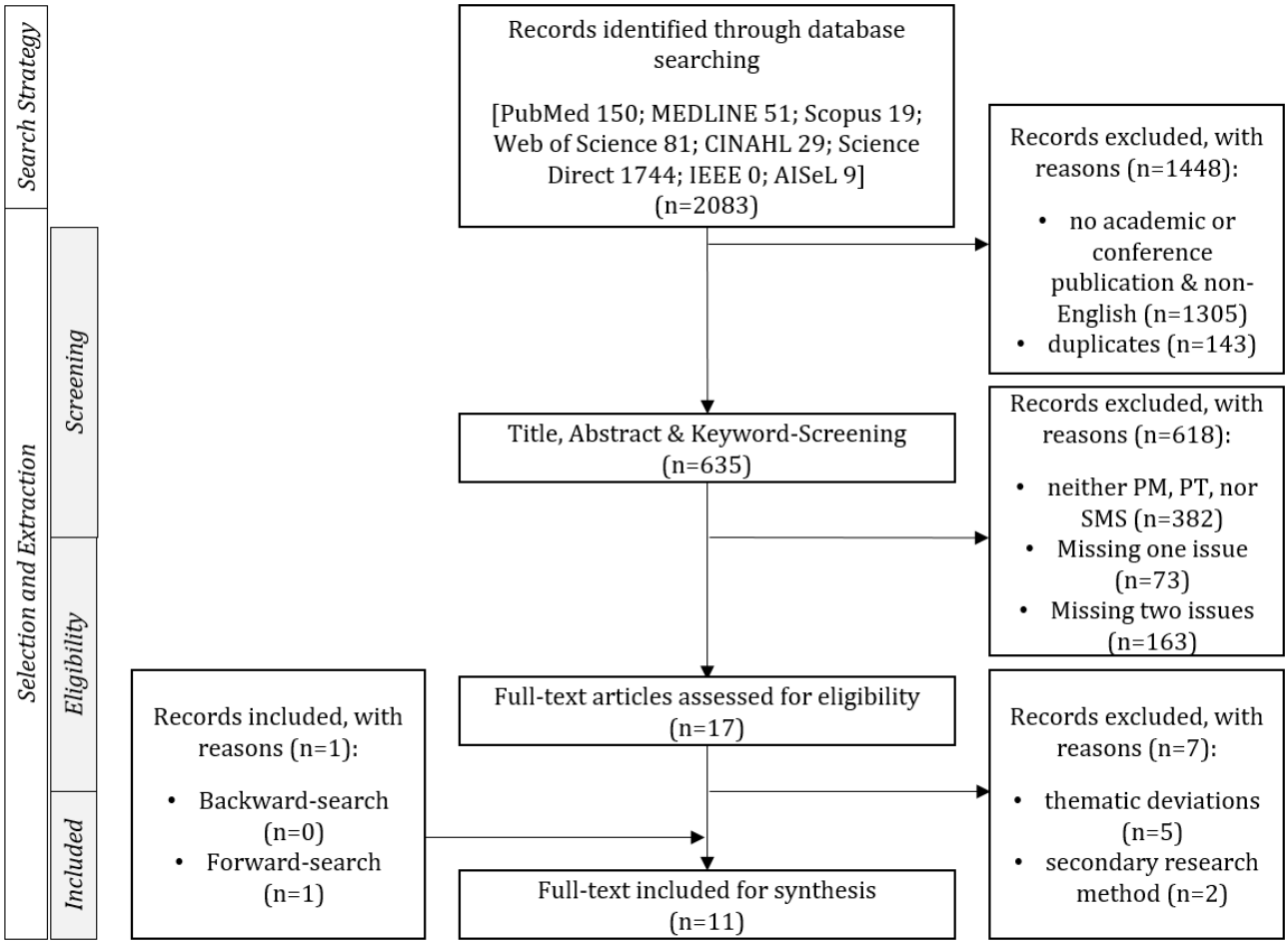
The goal of this literature review was to provide comprehensive insight into postoperative management via SMS. The review should (1) provide information on the state of research, (2) establish a basic understanding of SMS-based pain management, and (3) provide deep insight into the feasibility of these management strategies. To ensure completeness and transparency, a systematic literature review process was followed in all stages of the study. The methodology is mainly based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [28] and a guide for systematic literature reviews in information systems research [29]. A detailed and well-structured protocol was initially drafted, which fully defines the procedures to ensure validity and accuracy [28,30]. The protocol can be found in Multimedia Appendix 1.

Search Strategy

In consideration of the research question, a search string was designed using Boolean operators (AND, OR) for the selection of relevant literature. Various keywords connected the decisive aspects regarding text messages, postoperative care, and pain management. The conclusive search string is documented in the protocol (Multimedia Appendix 1) and is composed in the detailed form of “text messag*” OR “short messag*” OR “sms” AND “postoperative care*” OR “postoperative care*” OR “surge*” OR “surgic*” OR “operat*” AND “pain” OR “medicat*” OR “opioid*” OR “analgesic.” The searches were performed in three databases in the medical field (PubMed, Medline, and CINAHL), three interdisciplinary databases (Web of Science, Scopus, and Science Direct), and two databases

covering the field of computer science (IEEE and AISeL) (Figure 1). The search string was adapted to the specific characteristics and requirements of the respective database. The search was performed between March and April 2019.

Figure 1. Systematic literature review flowchart. PM: pain management; PT: postoperative treatment.



Selection and Extraction

The initial search in the eight databases resulted in 2083 matches. The screening process narrowed down this result by applying a variety of exclusion criteria. Initially, only articles that were published in academic journals or conferences were considered. Further, all non-English articles were excluded. By applying these two criteria, 1305 articles were already excluded. Next, all duplicates were identified by DOI or title alignment and deleted for the next steps (n=143). The remaining 635 articles were screened by title, abstract, and keywords for their relevance to the research question. Accordingly, the papers needed to describe and analyze pain management, and to clearly emphasize postoperative care and SMS. Articles that interpreted abbreviations such as SMS differently or focused on issues not relevant to the review were further excluded. During this screening process, a total of 382 articles could be identified as irrelevant, as these articles neither focused on pain management, postoperative treatment, nor SMS. Only two of these three issues were addressed in 73 articles and one of the three was addressed in 163 articles. As a result of this process, 618 articles were classified as irrelevant, leaving a total of 17 articles eligible for further review.

Full-text screening was performed to comprehensively analyze the remaining 17 articles, and an additional 7 articles were excluded due to thematic deviations or the secondary research method. Based on the remaining 10 articles, a forward and backward search was executed. The backward search revealed no new articles. The forward search resulted in a total of 64 matches for the eight databases. After applying the exclusion criteria, one additional article could be included for analysis, leaving 11 articles in the final pool for review (Figure 1) [28,31].

Results

Characteristics of Selected Studies

Of the 635 articles analyzed in the screening process, 554 (87.2%) were published in the last 5 years, since 2014. This reflects the increasing relevance of digitalized postoperative pain management in recent years. Correspondingly, this is apparent in the final pool of 11 articles, with 45% (n=5) published in 2018 and 36% (n=4) published in 2019. Three articles were published in the journal *Telemedicine and e-Health* [32-34], representing the interface between health and informatics. The remaining eight articles were published in various health-related journals. The average 2018 impact factor of the journals was 2.51, with the best-ranked journal at 6.03.

Ten of these 11 (91%) studies adopted a prospective design to investigate postoperative pain management in conjunction with

SMS. Patients were grouped as cohorts, and investigated for pain and similar outcomes. One study was a nonblinded randomized control trial [35]. The main research aim of 8 of the 11 (73%) studies was to determine the intervention's feasibility (Table 1).

Table 1. Characteristics of the 11 studies.

Reference	Study design	Research aim	Automation	Age (years), mean (SD)	Surgical procedure
Anthony et al [32]	Prospective multicenter cohort	Feasibility	yes	49.6 (13.7)	Hand surgery
Anthony et al [33]	Prospective cohort	Feasibility	yes	46.0 (22.0)	Lower extremity fracture
Booth et al [36]	Prospective cohort	Feasibility	n/a[a]	30.7 (5.5)	Caesarean section
Brix et al [35]	Nonblinded randomized control trial	Feasibility	yes	47.5 (16.5)	Knee arthroplasty
Carrier et al [37]	Prospective multicenter cohort	Validation	yes	57.0 (n/a)	Colorectal surgery
Chen et al [38]	Prospective cohort	Feasibility	yes	8.5 (n/a)	Tonsillectomy
Day et al [39]	Prospective cohort	Feasibility	yes	n/a	Total hip or knee arthroplasty
Nelson et al [40]	Prospective cohort	Feasibility	yes	6.1 (2.1)	Humeral fractures
Newton and Sulman [41]	Prospective cohort	Feasibility	n/a	n/a	Tonsillectomy
Premkumar et al [42]	Prospective cohort	Validation	yes	59.4 (10.9)	Total hip or knee arthroplasty
Yahanda et al [34]	Prospective cohort	Validation	yes	n/a	Total hip or knee arthroplasty

[a] n/a: not available; the article did not provide corresponding information.

Characteristics of Study Populations

Overall, 4195 patients were supported by a pain management system tailored to the surgery performed and to the patients' characteristics such as age or physical condition. The number of study participants ranged from 21 to 3049, with a mean of 381 participants and a median of 85. To select these patients, the researchers applied various selection criteria. For instance, 9 of the 11 (82%) research teams excluded patients without a mobile phone, 6 (55%) excluded patients who could not communicate via SMS, and 5 (45%) excluded patients with a language barrier. At the beginning of the intervention, the sex and age, and other demographic data of all participants were determined in 8 of 11 (73%) studies. The percentage of female participants ranged between 33% and 100%. The intervention was directed at adult patients with a mean age between 30.7 and 59.4 years in 6 of 8 (75%) studies. The other 2 studies analyzed pain management in children between 6.1 and 8.2 years of age (Table 1). The educational level and work status were determined in 4 of these 8 (50%) studies. The percentage of participants with an educational level above a bachelor's degree varied between 23% and 43%, and the proportion of participants with full-time employment ranged between 40% and 75%. The race, BMI, and surgical history of the patients were identified in 3 of the 8 (38%) studies. Six of the 11 (55%) studies carried out a software demonstration.

Intervention Characteristics of Selected Studies

Nine of the 11 studies used automated pain management systems for their interventions and the other 2 studies did not provide any relevant information on this aspect [36,41]. For 7 of the 11 (64%) studies, the primary reason for sending text messages was to monitor patients' postoperative pain. The remaining 4 (36%) studies intended to determine adherence to the pain treatment, which was combined with a teaching purpose in 2 of these studies and with a monitoring assignment in the other 2 studies. The content of the text messages and other monitoring aspects was directly linked to the various surgeries performed (Table 2). Four (36%) studies examined different outcomes after knee or hip surgery, two each following tonsillectomy or fracture (18%), and one each after a C-section, hand surgery, or colorectal surgery (9%) (Table 1).

Table 2. Overview of the interventions.

Study	Message purpose	Time (postoperative days)	Opioids	Pain scale	Alerts	Reminders
Anthony et al [32]	Monitoring	7	yes[a]	0-10	no	yes
Anthony et al [33]	Monitoring	14	yes	0-10	no	no
Booth et al [36]	Monitoring	60	yes	0-10	yes	no
Brix et al [35]	Adherence/monitoring	4	no	0-10	yes	no
Carrier et al [37]	Monitoring	7	yes[a]	0-10	no	yes
Chen et al [38]	Monitoring	14	yes[a]	0-10	no	yes
Day et al [39]	Education/adherence	14	yes[a]	no	yes	no
Nelson et al [40]	Monitoring	21	yes	0-10	no	no
Newton and Sulman [41]	Education/adherence	9	yes[a]	no	yes	no
Premkumar et al [42]	Monitoring	42	yes[a]	0-10	no	no
Yahanda et al [34]	Adherence/monitoring	15	yes[a]	0-9	no	yes

[a]No specific information about the type of opioid used given.

Process of the Interventions

Whether the goal of the intervention was for patient monitoring or to analyze the adherence to pain management, the studies defined different timeframes. With 60 intervention days, one study was distinctly longer than the others [36] and can therefore be considered an outlier. The average timeframe of the remaining 10 interventions was 14.7 days for postoperative care. The individual outcomes were determined by various message blocks consisting of several SMS messages sent on different postoperative days. Eight of the 11 (73%) interventions transmitted at least one of these blocks daily, 1 study sent a message every second day [37], and 2 studies each sent messages at a certain time interval [40,41]. The first intervention informed patients daily in the first week, and then on the 10th, 14th, and 21st postoperative day [40]. The other intervention sent messages daily for the first 3 postoperative days, and then on the 5th, followed by the 7th to 9th postoperative days daily [41]. Within these time intervals, six interventions measured the outcome with one message block, usually in the morning. Two interventions sent one message block each morning and evening to the patients [40,42]. Three interventions sent an additional block at 12 PM [32,33,35]. Each of the blocks consisted of at least one question sent by SMS. In 7 of the 11 (64%) interventions, one block was subdivided into at least three SMS.

For 7 of the 11 (64%) studies, the message blocks' primary outcome was the monitoring of the postsurgical pain level of the investigated patients. However, 3 of the 11 (27%) studies focused primarily on adherence to the treatment [34,35,41], and 1 (9%) study focused on the satisfaction level of the patients [39]. Further secondary results included drug intake, patient satisfaction, or the number of alert messages, among others. In total, 9 of the 11 (81%) studies measured current pain levels, either as a primary or as a secondary outcome. For this purpose, 8 of these studies used a pain scale from 0 to 10, and one used a scale from 0 to 9 [34]. Response 0 always reflects little or no pain, and 9 or 10 indicates the most severe pain. Three interventions focused on postoperative pain management for children, thereby involving legal guardians in the process [38,40,41]. For this monitoring, two studies utilized Wong-Baker Faces [38,40], which is a valid and effective method of assessing pain in children [43]. Postoperatively, patient groups received opioids in 10 interventions. Percocet [33], oxycodone [33,40], morphine [33,36], or fentanyl [36] was prescribed depending on the intensity of pain and severity of the procedure (Table 2).

Four out of the 11 (36%) interventions sent additional reminders to patients, and 4 interventions (36%) also sent alerts to patients and physicians (Table 2). Patients were reminded when they missed answering questions for several days [36], and one intervention provided three daily reminders regarding medication intake [35]. Alerts were triggered whenever various thresholds or schedules were exceeded, or when communication with the system occurred at an unscheduled time [32,34,37,38]. Once alerted, health care professionals organized the appropriate actions, initially by contacting the patients directly. Depending on these alerts, reminders, current symptoms, and responses to the respective message blocks, the pain management could be adjusted. For instance, alerts could lead to a change in medication or a follow-up examination.

Intervention Results of Selected Studies

All studies identified a positive effect of SMS on pain management, thereby indirectly providing various recommendations for action. First, 7 interventions measured response rates, each with concrete results. Overall, between 8 and approximately 400 messages were sent to patients or legal guardians. With an average response rate of approximately 89.5% (SD 3.8%), the reliability of SMS as a communication and monitoring tool is evident, especially in comparison to conventional communication methods (Table 3). One study referred to two interventions where the response rate was 63%

for telephone calls and 72% for a mobile app [36]. The response rate via SMS was significantly higher than that for email interventions (66.63%, $P<.001$) [42]. However, the researchers were unable to establish an association between response rate and age, level of education, and working status of the patients.

Further, the response rate was confirmed to decrease steadily over the intervention duration, and the majority of unanswered messages occurred within the last postoperative days [33,37,38]. The highest pain levels were measured in the first postoperative days [32,33,38,40] and then decreased daily to clinically unimportant levels [38,40]. One study showed that postoperative opioid use had a strong positive correlation with the reported pain ($r=0.972$, $P<.001$) [40]. All interventions that assessed medication intake confirmed this trend [32,33]. Six studies were verified and approved by respective institutional review boards (IRBs). Utilizing SMS for data collection was further deemed to be compliant with the Health Insurance Portability and Accountability Act (HIPAA) in three studies [32,33,40] (Table 3).

Table 3. Summary of systematic literature review results.

Reference	Outcome	Results	SMS messages sent, N	Response rate (%)	Compliance
Anthony et al [32]	Effective	Highest pain level within first 48 h; average use of 15.9 prescription opioids	~19	88.3	IRB[a]/HIPAA[b]
Anthony et al [33]	Effective	Response rate, pain, and medication intake decline over time	~22	87.5	IRB/HIPAA
Booth et al [36]	Positive impact	Rate especially powerful compared to traditional methods	~400	82.0	IRB
Brix et al [35]	Positive and efficient	Nonsignificant trend for better adherence	~8	n/a[c]	n/a
Carrier et al [37]	Positive impact	Intervention led to earlier detection	16	89.5	n/a
Chen et al [38]	Positive impact	Real-time monitoring possible	14	88.0	n/a
Day et al [39]	Positive impact	High satisfaction rate, high adherence and acceptance among patients	~18	n/a	IRB
Nelson et al [40]	Positive impact	Less medication intake, pain decreased daily	~20	88.4	IRB/HIPAA
Newton and Sulman [41]	Positive impact	Improved adherence and communication quality, less anxiety, positive educational effect	12	n/a	n/a
Premkumar et al [42]	Positive impact	Real-time, highly accepted, and available data collection method	~80	96.1	n/a
Yahanda et al [34]	Positive impact	Improved adherence and satisfaction	~18	n/a	IRB

[a]IRB: institutional review board.

[b]HIPAA: Health Insurance Portability and Accountability Act.

[c]n/a: not available; the article does not provide corresponding information.

Discussion

Principal Results

SMS-based pain management is highly applicable and efficient for postoperative communication between health care professionals and patients or legal guardians. Furthermore, alarms and reminders via SMS can improve and maintain communication, while supporting patients or their legal guardians. This support function is desirable for effective pain management [44]. Three of the 11 (27%) interventions even preoperatively communicated with patients [34,39] or their legal guardians [41]. This preoperative communication via SMS yields equally positive results as postoperative pain management; various studies confirmed this conclusion. Patients appear to be satisfied with preoperative preparation before treatment [45], while the legal guardians considered an automated SMS system as a beneficial support system [46]. One study utilized days 7, 4, 2, and 1 [39], whereas another used days 14, 4, 2, and 1 before surgery [41] for the intervention. Both studies relied on SMS for providing additional information regarding the surgery and specific process steps. Moreover, the text messages functioned as reminders to adhere to schedules and specific requirements or prerequisites of the intervention. The third study used two preoperative SMS messages to ensure that the prescribed medication was purchased and used correctly by the patient, starting 6 days before surgery [34]. Accordingly, 10 interventions averaged 17.4 days for pre- and postoperative care, and 14.7 days for purely postoperative care.

SMS technology was associated with positive results for all studies. The ubiquity of SMS makes it a cost-effective and straightforward method for pain management that is valid and less intrusive [35,38,40,42]. In addition to the very high response rate, the patients also responded quickly. One study determined an average response time of fewer than 12 minutes [37], further demonstrating the increased acceptance and utilization of SMS. This could enable the critical drivers of the opioid epidemic, such as inadequate medical education and guidelines, to be addressed directly, and most likely very effectively and

efficiently. Nine interventions successfully monitored pharmaceutical opioids, and individually adjusted the pain management according to pain perception and response to the respective message blocks. The SMS systems improved adherence to pain management, and one study even reported less medication intake overall [40]. Patients could be reliably contacted, facilitating valuable information, education, and further questions. Therefore, an implementation of SMS-based pain management could combat the opioid crisis. Toward this end, process automation could be a crucial aspect. Automated SMS systems for pain management enable more robust data collection without consuming limited health resources, especially regarding personnel.

Furthermore, the use of SMS can prevent potential bias, and ensure the consistency and timeliness of messaging to patients [32,33,35,37,42]. Automatic postoperative communication has already been shown to reduce opioid intake in patients with orthopedic trauma [47]. One study also emphasized that automated alerts have enabled the more efficient and effective detection of postoperative complications [37]. These included, among others, pain above a certain level, no responses, missing acquisition of medications, and specific symptoms [34,37]. The information elicited through SMS had a net benefit in fewer telephone calls, saving time and personnel costs [39]. Generally, the constant messages and communication led to a positive patient experience. SMS improved the patients' understanding and responsibility, and reduced their anxiety regarding the operation [39,41].

SMS-based pain management allows for simple pre- and postoperative extensions such as easily integrable and more specific questions about the operation and possible symptoms. In addition, reminders and alerts can be triggered automatically by SMS systems. Nevertheless, the response rate was confirmed to decrease steadily over the intervention duration, and the majority of unanswered SMS messages occurred within the last postoperative days [33,37,38]. Therefore, the extensions should be limited and the pain management period should be as short as possible, depending on the operation. The development of pain intensity and medication intake during each intervention supports this assumption. In conclusion, depending on the operation and the associated pain intensity, a monitoring timeframe and a medication schedule should be defined.

Limitations

This study is subject to various limitations. First, only a basic quality analysis of the identified studies was applied. A more detailed analysis could clarify whether the final pool is rigorous, relevant, and credible. However, demand for high-quality research approaches such as randomized control studies is identified, as only one study was a nonblinded randomized control trial [35]. Second, the utilized Boolean search string could be defined more precisely. For instance, the results indicated that preoperative pain management has a decisive influence on an intervention's success. Therefore, keywords such as "preoperative care*" or "presurgical care*" should be added. By enriching the search string, a more accurate result could be obtained.

Conclusions

This study provides a comprehensive review of the current status of the literature on postoperative pain management by SMS. SMS utilization as a communication channel appeared to be favorable and feasible in pain management in the postoperative phase. According to three studies, SMS also seems to be useful for preoperative pain management, especially for additional information on medication or schedules, or as reminders. SMS resulted in excellent patient response rates, better adherence to pain treatment, higher patient satisfaction, and less medication intake. Six SMS interventions were certified by the respective IRBs and three were HIPAA-compliant. This indicates that SMS is capable of meeting health care requirements and is suitable for a health care–specific application. All of these benefits could be leveraged to address the opioid epidemic directly, effectively, and efficiently. The ability to create efficient pain management via SMS ensures comprehensive monitoring and communication. Key drivers of the opioid epidemic, such as medication abuse, shifted patient expectations, inadequate medical education, or inadequate guidelines, could be adequately addressed. In conclusion, the research question could be confirmed: SMS is effective, very well-suited, and feasible for postoperative pain management.

Future research should extend this systematic literature review regarding preoperative pain management. Based on this, a pre- and postoperative communication model should be developed to address the opioid epidemic effectively. This model should be generally applicable and adaptable to the individual clinical situation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Systematic review protocol.
[PDF File (Adobe PDF File), 246 KB - mhealth_v9i6e20199_app1.pdf]

References

1. Fischer B, Keates A, Bühringer G, Reimer J, Rehm J. Non-medical use of prescription opioids and prescription opioid-related harms: why so markedly higher in North America compared to the rest of the world? Addiction 2014 Feb;109(2):177-181. [doi: 10.1111/add.12224] [Medline: 23692335]

2. UNODC. World Drug Report 2019: (Set of 5 Booklets). S.l. New York: United Nations; 2019:978-992.
3. Hedegaard H, Miniño AM, Warner M. Drug overdose deaths in the United States, 1999-2017. NCHS Data Brief 2018 Nov(329):1-8 [FREE Full text] [Medline: 30500323]
4. Centers for Disease Control and Prevention. 2018 Annual Surveillance Report of Drug-Related Risks and Outcomes - United States. 2018. URL: https://www.cdc.gov/drugoverdose/pdf/pubs/2018-cdc-drug-surveillance-report.pdf [accessed 2021-05-22]
5. Gostin LO, Hodge JG, Noe SA. Reframing the opioid epidemic as a national emergency. JAMA 2017 Oct 24;318(16):1539-1540. [doi: 10.1001/jama.2017.13358] [Medline: 28832871]
6. Madras BK. The President's Commission on combating drug addiction and the opioid crisis: origins and recommendations. Clin Pharmacol Ther 2018 Jun;103(6):943-945. [doi: 10.1002/cpt.1050] [Medline: 29570781]
7. Christie C, Baker C, Cooper R, Kennedy P, Madras B, Bondi P. The President's Commission on Combating Drug Addiction and the Opioid Crisis. Homeland Security Digital Library.: The Office of National Drug Control Policy; 2017. URL: https://www.doh.wa.gov/Portals/1/Documents/2300/2017/PresidentsCommissionOnCombatingDrugAddictionOpioidCrisis.pdf [accessed 2021-05-22]
8. Federation of State Medical Boards of the United States, Inc. Model policy for the use of controlled substances for the treatment of pain. J Pain Palliat Care Pharmacother 2005;19(2):73-78. [Medline: 16061467]
9. Volkow ND, McLellan TA, Cotto JH, Karithanom M, Weiss SRB. Characteristics of opioid prescriptions in 2009. JAMA 2011 Apr 06;305(13):1299-1301 [FREE Full text] [doi: 10.1001/jama.2011.401] [Medline: 21467282]
10. McIver JS. Seeking solutions to the opioid crisis. P T 2017 Jul;42(7):478 [FREE Full text] [Medline: 28674476]
11. Wiese HJC, Piercey RR, Clark CD. Changing prescribing behavior in the United States: moving upstream in opioid prescription education. Clin Pharmacol Ther 2018 Jun;103(6):982-989. [doi: 10.1002/cpt.1015] [Medline: 29315508]
12. Bohnert ASB, Valenstein M, Bair MJ, Ganoczy D, McCarthy JF, Ilgen MA, et al. Association between opioid prescribing patterns and opioid overdose-related deaths. JAMA 2011 Apr 06;305(13):1315-1321. [doi: 10.1001/jama.2011.370] [Medline: 21467284]
13. Henry SG, Matthias MS. Patient-clinician communication about pain: a conceptual model and narrative review. Pain Med 2018 Nov 01;19(11):2154-2165 [FREE Full text] [doi: 10.1093/pm/pny003] [Medline: 29401356]
14. Bensing J, van Dulmen S, Kallerup H, Visser A, Borrell F, Finset A, et al. The European Association for Communication in Healthcare. Patient Educ Counsel 2001 Apr;43(1):1-4. [doi: 10.1016/s0738-3991(01)00125-2]
15. McLafferty RB, Williams RG, Lambert AD, Dunnington GL. Surgeon communication behaviors that lead patients to not recommend the surgeon to family members or friends: Analysis and impact. Surgery 2006 Oct;140(4):616-22; discussion 622. [doi: 10.1016/j.surg.2006.06.021] [Medline: 17011909]
16. Coiera E. Communication systems in healthcare. Clin Biochem Rev 2006 May;27(2):89-98 [FREE Full text] [Medline: 17077879]
17. Apfelbaum JL, Chen C, Mehta SS, Gan TJ. Postoperative pain experience: results from a national survey suggest postoperative pain continues to be undermanaged. Anesth Analg 2003 Aug;97(2):534-40, table of contents. [doi: 10.1213/01.ane.0000068822.10113.9e] [Medline: 12873949]
18. Haggerty E. Healthcare and digital transformation. Netw Sec 2017 Aug;2017(8):7-11. [doi: 10.1016/s1353-4858(17)30081-8]
19. Weiner JP. Doctor-patient communication in the e-health era. Isr J Health Policy Res 2012 Aug 28;1(1):33 [FREE Full text] [doi: 10.1186/2045-4015-1-33] [Medline: 22929000]
20. Gagnon M, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277 [FREE Full text] [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]
21. Bergmo TS, Kummervold PE, Gammon D, Dahl LB. Electronic patient-provider communication: will it offset office visits and telephone consultations in primary care? Int J Med Inform 2005 Sep;74(9):705-710. [doi: 10.1016/j.ijmedinf.2005.06.002] [Medline: 16095961]
22. Lu K, Marino NE, Russell D, Singareddy A, Zhang D, Hardi A, et al. Use of short message service and smartphone applications in the management of surgical patients: a systematic review. Telemed J E Health 2018 Jun;24(6):406-414. [doi: 10.1089/tmj.2017.0123] [Medline: 29111887]
23. Tseng C, Jiang J, Lee R, Lu F, Ouyang C, Chen Y, et al. Feasibility study on application of GSM–SMS technology to field data acquisition. Comput Electron Agric 2006 Aug;53(1):45-59. [doi: 10.1016/j.compag.2006.03.005]
24. Church K, Oliveira RD. What's up with Whatsapp? New York, NY: ACM Press; 2013 Presented at: 15th international conference on Human-computer interaction with mobile devices and services; August 2013; Munich p. 13. [doi: 10.1145/2493190.2493225]
25. Mobile fact sheet 2017. Pew Research Center. URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2020-12-07]
26. Smith A. U.S. Smartphone Use in 2015. Pew Research Center Internet & Technology. URL: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/ [accessed 2020-12-07]
27. Miloh T, Annunziato R, Arnon R, Warshaw J, Parkar S, Suchy FJ, et al. Improved adherence and outcomes for pediatric liver transplant recipients by using text messaging. Pediatrics 2009 Nov;124(5):e844-e850. [doi: 10.1542/peds.2009-0415] [Medline: 19822583]

28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
29. Okoli C, Schabram K. A guide to conducting a systematic literature review of information systems research. SSRN Journal 2010:1-49. [doi: 10.2139/ssrn.1954824]
30. 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. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
31. Thuan NH, Antunes P, Johnstone D. Factors influencing the decision to crowdsource: A systematic literature review. Inf Syst Front 2015 Jun 26;18(1):47-68. [doi: 10.1007/s10796-015-9578-x]
32. Anthony CA, Lawler EA, Ward CM, Lin IC, Shah AS. Use of an automated mobile phone messaging robot in postoperative patient monitoring. Telemed J E Health 2018 Jan;24(1):61-66. [doi: 10.1089/tmj.2017.0055] [Medline: 28622079]
33. Anthony CA, Volkmar A, Shah AS, Willey M, Karam M, Marsh JL. Communication with orthopedic trauma patients via an automated mobile phone messaging robot. Telemed J E Health 2018 Jul;24(7):504-509. [doi: 10.1089/tmj.2017.0188] [Medline: 29261036]
34. Yahanda AT, Marino NE, Barron J, Concepcion A, St John T, Lu K, et al. Patient engagement and cost savings achieved by automated telemonitoring systems designed to prevent and identify surgical site infections after joint replacement. Telemed J E Health 2019 Feb;25(2):143-151 [FREE Full text] [doi: 10.1089/tmj.2017.0325] [Medline: 30192209]
35. Brix LD, Bjørnholdt KT, Thillemann TM, Nikolajsen L. The effect of text messaging on medication adherence after outpatient knee arthroscopy: a randomized controlled trial. J Perianesth Nurs 2019 Aug;34(4):710-716. [doi: 10.1016/j.jopan.2018.11.011] [Medline: 30852173]
36. Booth JL, Sharpe EE, Houle TT, Harris L, Curry RS, Aschenbrenner CA, et al. Patterns of recovery from pain after cesarean delivery. Pain 2018 Oct;159(10):2088-2096 [FREE Full text] [doi: 10.1097/j.pain.0000000000001313] [Medline: 29905650]
37. Carrier G, Cotte E, Beyer-Berjot L, Faucheron JL, Joris J, Slim K, Groupe Francophone de Réhabilitation Améliorée après Chirurgie (GRACE). Post-discharge follow-up using text messaging within an enhanced recovery program after colorectal surgery. J Visc Surg 2016 Aug;153(4):249-252. [doi: 10.1016/j.jviscsurg.2016.05.016] [Medline: 27423211]
38. Chen Y, Chin M, Greenberg S, Johnstone C, McGuinness J. Post-tonsillectomy pain in 24 children - utilising short message service (SMS) to assess postoperative outcomes. Clin Otolaryngol 2012 Oct;37(5):412-414. [doi: 10.1111/j.1749-4486.2012.02521.x] [Medline: 23164270]
39. Day MA, Anthony CA, Bedard NA, Glass NA, Clark CR, Callaghan JJ, et al. Increasing perioperative communication with automated mobile phone messaging in total joint arthroplasty. J Arthroplasty 2018 Jan;33(1):19-24. [doi: 10.1016/j.arth.2017.08.046] [Medline: 29017803]
40. Nelson SE, Adams AJ, Buczek MJ, Anthony CA, Shah AS. Postoperative pain and opioid use in children with supracondylar humeral fractures: balancing analgesia and opioid stewardship. J Bone Joint Surg Am 2019 Jan 16;101(2):119-126. [doi: 10.2106/JBJS.18.00657] [Medline: 30653041]
41. Newton L, Sulman C. Use of text messaging to improve patient experience and communication with pediatric tonsillectomy patients. Int J Pediatr Otorhinolaryngol 2018 Oct;113:213-217. [doi: 10.1016/j.ijporl.2018.07.048] [Medline: 30173988]
42. Premkumar A, Lovecchio FC, Stepan JG, Kahlenberg CA, Blevins JL, Albert TJ, et al. A novel mobile phone text messaging platform improves collection of patient-reported post-operative pain and opioid use following orthopedic surgery. HSS J 2019 Feb;15(1):37-41 [FREE Full text] [doi: 10.1007/s11420-018-9635-3] [Medline: 30863231]
43. Garra G, Singer AJ, Taira BR, Chohan J, Cardoz H, Chisena E, et al. Validation of the Wong-Baker FACES Pain Rating Scale in pediatric emergency department patients. Acad Emerg Med 2010 Jan;17(1):50-54. [doi: 10.1111/j.1553-2712.2009.00620.x] [Medline: 20003121]
44. Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, et al. Commercially available smartphone apps to support postoperative pain self-management: scoping review. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e162 [FREE Full text] [doi: 10.2196/mhealth.8230] [Medline: 29061558]
45. Walter BM, Klare P, Neu B, Schmid RM, von Delius S. Development and testing of an automated 4-day text messaging guidance as an aid for improving colonoscopy preparation. JMIR Mhealth Uhealth 2016 Jun 21;4(2):e75 [FREE Full text] [doi: 10.2196/mhealth.5289] [Medline: 27329204]
46. Farias N, Rose-Davis B, Hong P, Wozney L. An automated text messaging system (Tonsil-Text-To-Me) to improve tonsillectomy perioperative experience: exploratory qualitative usability and feasibility study. JMIR Perioper Med 2020 Jan 15;3(1):e14601 [FREE Full text] [doi: 10.2196/14601] [Medline: 33393931]
47. Anthony CA, Rojas EO, Keffala V, Glass NA, Shah AS, Miller BJ, et al. Acceptance and commitment therapy delivered via a mobile phone messaging robot to decrease postoperative opioid use in patients with orthopedic trauma: randomized controlled trial. J Med Internet Res 2020 Jul 29;22(7):e17750 [FREE Full text] [doi: 10.2196/17750] [Medline: 32723723]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act

IRB: institutional review board

Edited by G Eysenbach, Q Zeng; submitted 13.05.20; peer-reviewed by L Brix, J Swan; comments to author 28.08.20; revised version received 19.12.20; accepted 13.04.21; published 16.06.21.

Please cite as:
Buck C, Keweloh C, Bouras A, Simoes EJ
Efficacy of Short Message Service Text Messaging Interventions for Postoperative Pain Management: Systematic Review
JMIR Mhealth Uhealth 2021;9(6):e20199
URL: https://mhealth.jmir.org/2021/6/e20199
doi:10.2196/20199
PMID:34132646

Original Paper

Motivating Adherence to Exercise Plans Through a Personalized Mobile Health App: Enhanced Action Design Research Approach

Ruo-Ting Sun[1], MS; Wencui Han[1], PhD; Hsin-Lu Chang[2], PhD; Michael J Shaw[1], PhD

[1]Department of Business Administration, University of Illinois, Urbana-Champaign, Champaign, IL, United States

[2]Department of Management Information Systems, National Chengchi University, Taipei, Taiwan

Corresponding Author:
Hsin-Lu Chang, PhD
Department of Management Information Systems
National Chengchi University
64, Sec 2, Chinan Rd,
Taipei, 116
Taiwan
Phone: 886 2 29393091 ext 81214
Fax: 886 2 29351552
Email: hlchang@nccu.edu.tw

Abstract

Background: Physical inactivity is a global issue that affects people's health and productivity. With the advancement of mobile technologies, many apps have been developed to facilitate health self-management. However, few studies have examined the effectiveness of these mobile health (mHealth) apps in motivating exercise adherence.

Objective: This study aims to demonstrate the enhanced action design research (ADR) process and improve the design of mHealth apps for exercise self-management. Specifically, we investigate whether sending motivational messages improves adherence to exercise plans, whether the motivational effect is affected by personality, the impact of message type and repetition, and the process of involving a field experiment in the design process and learning new design principles from the results.

Methods: This formative research was conducted by proposing an enhanced ADR process. We incorporated a field experiment into the process to iteratively refine and evaluate the design until it converges into a final mHealth app. We used the Apple ResearchKit to develop the mHealth app and promoted it via trainers at their gyms. We targeted users who used the app for at least two months. Participants were randomly assigned to 1 of the 12 groups in a 2×3×2 factorial design and remained blinded to the assigned intervention. The groups were defined based on personality type (thinking or feeling), message type (emotional, logical, or none), and repetition (none or once). Participants with different personality types received tailored and repeated messages. Finally, we used the self-reported completion rate to measure participants' adherence level to exercise plans. By analyzing users' usage patterns, we could verify, correct, and enhance the mHealth app design principles.

Results: In total, 160 users downloaded the app, and 89 active participants remained during the 2-month period. The results suggest a significant main effect of personality type and repetition and a significant interaction effect between personality type and repetition. The adherence rate of people with feeling personality types was 18.15% higher than that of people with thinking types. Emotional messages were more effective than logical messages in motivating exercise adherence. Although people received repeated messages, they were more likely to adhere to exercise plans. With repeated reminders, the adherence rates of people with thinking personality types were significantly improved by 27.34% (P<.001).

Conclusions: This study contributes to the literature on mHealth apps. By incorporating a field experiment into the ADR process, we demonstrate the benefit of combining design science and field experiments. This study also contributes to the research on mHealth apps. The principles learned from this study can be applied to improve the effectiveness of mHealth apps. The app design can be considered a foundation for the development of more advanced apps for specific diseases, such as diabetes and asthma, in future research.

(JMIR Mhealth Uhealth 2021;9(6):e19941) doi:10.2196/19941

KEYWORDS

adherence; mobile health; motivation; personality; MBTI; action design research; mobile phone

Introduction

Background

In modern society, many people live a fast-paced, high-stress lifestyle and do not engage in regular physical activity. A study showed that 80% of adults in America lack physical exercise [1]. A lack of exercise is the main cause of most chronic diseases [2] and the leading cause of death worldwide [3,4]. Research also suggests that regular exercise improves the mood and self-reported performance of white-collar workers [5]. With the advancement of mobile and wearable technologies, many mobile health (mHealth) apps have been developed to enhance exercise adherence and facilitate a healthy lifestyle. mHealth apps are defined as "mobile applications that assist consumers in self-management of overall wellness, disease prevention, and disease management" [6]. The mHealth app market has been steadily growing over the past few years. According to an industry report [7], more than 318,000 mHealth apps are available in the top app stores worldwide, and more than 200 apps are added daily.

Despite the popularity of mHealth apps, questions remain regarding their sustained effectiveness. According to a recent survey [8], 87% of patients used mHealth apps, but more than one-quarter of them stopped using these apps because of ineffectiveness in helping achieve their health goals or yielding tangible results. Prior research has also noted that many mHealth apps lack theory-based motivational techniques, rendering it difficult for these apps to sustain over the long term [9]. Moreover, mHealth apps are often designed and treated as *black boxes*; the design is not evidence-based [10,11]. These findings suggest that mHealth app developers must design tools to be more engaging for users. However, few studies have examined the quality of mHealth apps in terms of their effectiveness in motivating health-related behaviors or the design principles of these apps from the user perspective [12-14].

Prior Work

Overview

Previous research discussed the optimal means of designing and evaluating persuasive systems and suggested a set of persuasive principles to design and evaluate these systems [15]. Persuasive principles are specific design techniques such as providing reminders, tailored and personalized information, *social support*, or suggestions. By including persuasive principles, mHealth apps can more effectively motivate users to adhere to their plans [16]. For example, a study conducted by Middelweerd et al [17] reviewed 64 of 41,246 mHealth apps available on Google Play and iTunes stores and suggested that techniques such as self-monitoring, receiving feedback on performance, and setting goals were most frequently used for persuasion. A review of mHealth apps that applied persuasive technology to improve physical exercise indicated that users were persuaded and became more involved in disease control or health management when persuasive design principles were applied [18]. Informed by the literature, in this study, we apply 3 persuasive design principles when designing an app and examined their effectiveness in motivating exercise. These principles included (1) tailored communication, (2) motivational messages, and (3) repetition of messages.

Tailored Communication

Past literature has suggested that the impact of health communication is generally enhanced when it is tailored to a specific individual [19]. However, the principles of audience segmentation are far less discussed in health care than in advertising, marketing, and social marketing [20,21]. In early work in this field, segmentation was often based on demographic differences. For example, different self-help guides for smoking cessation have been designed for blue-collar and minority smokers [22], African Americans [23], older smokers [24], pregnant women [25], and women with young children [26]. However, few studies have investigated the effectiveness of health communication by using *personality type* as a central tailoring variable. In this study, we followed the psychological type theory [27] and the Myers-Briggs Type Indicator (MBTI) test [28] to specifically consider 2 personality types, namely, the thinking type and the feeling type. The psychological type theory suggests that a person's seemingly random behavior is based on their inner preferences regarding perceiving and organizing information to form conclusions. The MBTI is a self-report questionnaire that can be used to assess people's psychological preferences in perceiving the world and making decisions. The MBTI is a well-adopted method used to quantify psychological types and is widely used in empirical studies to measure people's decision-making behavior [29-33]. The MBTI assesses personality types by considering a person's preference based on the following 4 pairs of psychological types: extraversion and introversion assess how people direct their energy either outwardly toward people and activities or inwardly toward thoughts and ideas; sensing and intuition refers to two ways of gathering information and understanding situations; thinking and feeling are two ways in which people organize and structure information and draw conclusions; and finally, judging and perceiving describes how people prefer to live their outer life. By adapting the MBTI, we were interested in investigating the impact of personality on the effectiveness of mHealth apps in promoting adherence to exercise plans. Therefore, thinking and feeling were selected as this pair represents the key dimension measuring how people organize and structure information and make decisions. People with a thinking personality type prefer applying analytical and logical principles to make objective decisions by following clear and consistent principles, whereas people with a feeling personality type may opt to make decisions by referencing their own and others' values and place more weight on personal concerns.

Motivational Messages

Several studies have investigated the effectiveness of motivational or persuasive messages in promoting targeted behaviors in the fields of psychology [34,35], marketing [36,37], public health [38,39], and health management [40-43]. Research has examined emotional versus rational messages [36,44,45], type of elaboration [37], positive versus negative messages, gain versus loss framing [39-41,46,47], and source credibility and likability [34]. The results of these studies consistently show that persuasive messages promote intended behaviors. In this

research, we were specifically interested in the effect of messages that generate positive emotions (*emotional* messages) compared with that of messages based on facts (*logical* messages). According to the Toulmin model of argumentation [48] and research conducted by Kim and Benbasat [49], a good argument with grounds (data), claims, and warrants leads to the highest level of trusting belief. We define logical messages as logical arguments consisting of a claim, data (ie, facts that support the claim), and backing (ie, data credibility). For example, "Don't forget to exercise today! Research shows that even one session of exercise will enhance your positive mood." The claim "Don't forget to exercise today," "even one session of exercise will enhance your positive mood" represents the data, and "research shows" provides the backing. In contrast, research has also shown that manipulating emotions accompanying a persuasive message affects the effectiveness of the message. People tend to adjust their beliefs to fit their emotions as people treat feelings as evidence [50]. Emotional stimuli can influence judgments without the judge's awareness of such stimuli [51]. In contrast to logical messages, emotional messages do not provide facts to support the claim but focus on triggering positive emotions (ie, claims plus positive emotion stimuli). For example, "It's time for your exercise! You are doing a fabulous job!"

Repetition of Messages

In addition to the message content, previous research has examined the impact of technical features of messages, such as message length [36], position [40,52], and repetition [37,53]. Among all the features, the most relevant and customizable feature in our context is repetition. The effect of repetition is referred to as the *mere repeated-exposure effect* in sociology, a psychological phenomenon involving people's tendency to develop a preference merely because they are repeatedly exposed to something [54]. A vast body of literature has shown that the repeated-exposure effect is a robust phenomenon demonstrated across cultures and diverse stimulus domains [55-60]. There are many applications of the mere repeated-exposure effect. In marketing, for example, many studies have tested the effects of advertisement repetition [61-63]. The repeated-exposure effect has also been studied in many other social and human decision-making contexts [62,64-66]. For example, in a study related to reminders using computer systems, Malone [67] suggested that displaying high-priority tasks more frequently is an effective reminder strategy. Therefore, we were interested in investigating the impact of message repetition on exercise adherence.

The Goal of This Study

This study's purpose was to improve the design of mHealth apps for exercise self-management by using an innovative research approach that combines field experiments with action design research (ADR). The literature examining adherence to exercise plans has defined and measured *adherence* as the percentage of completion of an exercise plan—often pertaining to plans agreed upon by patients and care providers in the case of medical studies [68,69]. We adopted this definition and measurement method by defining adherence to an exercise plan as the percentage of the exercise plan completed, and adherence was used as the dependent variable in this study. Notably, in this study, participants established their own plans. As they had very different health conditions and physical capabilities, it was not realistic for the researcher to create a universal plan. As the study participants likely had the desire to exercise, they set goals that they deemed beneficial to their health.

Specifically, our research objectives included (1) investigating whether sending motivational messages could improve adherence to exercise plans, (2) considering whether the motivational effect was impacted by personality, (3) testing message type (logical vs emotional) and repetition impact, and (4) exploring the possibility of involving a field experiment in the design process, learning new design principles from the results.

Methods

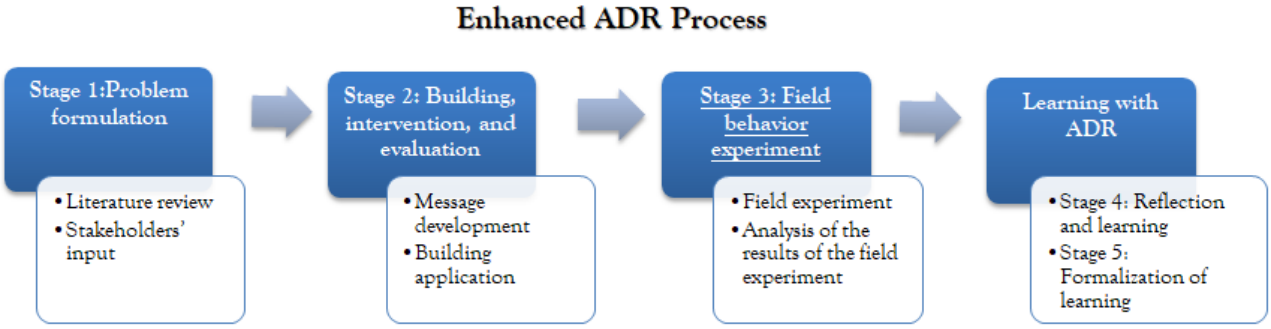
Overview

We followed an ADR approach to conduct this study. ADR, proposed by Sein et al [70], conceptualizes the research process as comprising the inseparable and inherently interwoven activities of building the information technology artifact, intervening in the organization, and concurrently evaluating the artifact. The process includes the following stages:

- Stage 1: problem formulation. In this stage, researchers and stakeholders determine the initial scope, decide the roles and scope of practitioner participation, and formulate the initial research questions.
- Stage 2: building, intervention, and evaluation. In this stage, based on the problem framing and theoretical premises adapted in stage 1, the research team builds the initial information technology artifact. The process should be performed as an iterative process in a targeted environment.
- Stage 3: reflection and learning. The reflection and learning stage proceeds conceptually from building a solution for a particular scenario to applying learning to a broader class of problems.
- Stage 4: formalization of learning. This stage generalizes the principles learned to a class of field problems.

To examine the effectiveness of an mHealth app in promoting adherence to an exercise plan, we enhanced the ADR process by adding a field experiment step. In this stage, we used artifacts to examine the specific design principles. By understanding how users interact with the artifact, we were able to link the behavior study with design research and then verify and correct the design principles. We demonstrate the research process in Figure 1 and detail the process in the following section.

Figure 1. Research process—enhanced action design research process. ADR: action design research.



Stage 1: Problem Formulation

On the basis of a literature review, review of web-based comments for mHealth apps, and focus groups with users, we identified mHealth apps as lacking theoretical guided design and found that few mHealth apps are focused on motivation, which is one of the main problems, along with the lack of evaluation of effectiveness, as detailed in the *Introduction* section. We identified app developers, users, and health care professionals (ie, gym trainers in our research settings) as stakeholders. We recruited 3 gym trainers and their trainees during the process of forming our initial research scope. While working with the researchers (among them, a participant was also the developer of the research app), all stakeholders provided input during the process. The gym trainers contributed their thoughts based on their prior experience with exercise management, that is, persuasive principles, such as daily progress reports of diet and exercise plans to supervise performance of trainees. We also interviewed trainees to learn about their prior experience using mHealth apps and summarized several necessary functions, such as event schedulers, alarm reminders, and progress reports. By understanding the mechanisms motivating the trainees, we and the developer discussed the executability of certain mHealth functionalities and how to maximize the effectiveness of motivational messages. The stakeholders can benefit from using our designed mHealth app as a tool to monitor and manage their exercise plans and learn persuasive principles. We defined the scope of the project as follows: develop an mHealth app to remind and motivate users to adhere to their exercise plans and examine the impact of personality and message characteristics on adherence. We intended to generate design principles for health self-management and other motivational applications.

Stage 2: Building, Intervention, and Evaluation

Message Selection

On the basis of the definition of logical and emotional messages detailed in the *Introduction* section, we selected and edited motivational messages from web-based forums. We first selected several popular sites where users post motivational messages related to exercise and fitness; then, messages that were clearly emotional or logical were manually selected. Next, we ensured that the messages selected were shorter than 128 characters to fit into a text message. We used lab testing to validate the message type classification, as described in the following sections.

Message Type Cross-validation

To classify the messages into emotional and logical types, we conducted 2 rounds of tests. Each round included 5 judges, who were students and faculty members in an information systems department and were not involved in the project. In the first round, the judges first read the definitions of logical and emotional messages and then classified each message as either an emotional or logical message. Any message that was not correctly classified by all the judges was removed from the message pool; we retained 30 logical and 30 emotional messages after the first round of classification. In the second round, we had another 5 judges to classify retained messages.

The interrater reliability measured by Cohen κ statistics was higher than 0.86 for any pair of judges (no more than 1 disagreement). After the 2 rounds of tests, we decided that the reliability of the classification of the messages had met the standards [71], so we used the messages in the following steps.

Pretest of the Message Impact

To pretest the impact of motivational messages, we recruited 100 college students to participate in pretesting. The participants took the MBTI test to determine whether they had a thinking or feeling personality type. Participants then read the motivational messages selected in the previous step and rated the perceived motivational level of each message on a 7-point Likert scale (strongly disagree to strongly agree with the effects of the motivational messages).

The average rating (mean 5.8, SD 1.6) of all messages suggested that participants found the messages motivational. Participants with a feeling personality type rated both emotional and logical messages higher than participants with a thinking personality type. The messages used in the final field experiment are presented in Multimedia Appendix 1.

Building the App

On the basis of the literature review and theoretical background detailed in the previous section, we envisioned an effective mHealth app that could set up alarm reminders for exercise (ie, message repetition) and send personalized motivational messages (ie, emotional or logical messages) based on users' personality types (ie, thinking type or feeling type). We used the Apple ResearchKit to develop an mHealth app named ActiveTrack. The idea behind the Apple ResearchKit is for scientists and drug developers to *build mobile or wearable apps* that suit their particular needs, whether for the collection of

research data, patient recruitment, or the collection of informed consent. A prototype of the app was built with the initial design and released to the Apple app store, and user activity data were saved in a web-based database. By observing app users' usage patterns, we identified points at which users were more likely to stop using the app. We also recruited users who provided feedback about the design and adjusted it based on these inputs. For example, in the initial design, users had to answer survey questions that collected their information, such as demographics, medical histories, and the MBTI questions, before entering the main page of the app. We observed that this caused most users to stop using the app. In total, 35 of the 79 inactive users dropped out at the beginning stage because of the overwhelming questions. Therefore, we reduced the number of questions and made the process shorter. We also made setting up reminders easier.

In the beta cycle, more users used ActiveTrack. Owing to the adjustments made, these users remained active for a longer time than the initial users. We hosted focus groups with some users and incorporated more suggestions into the design.

Design Artifact

The mHealth app (ActiveTrack) finalized for the field test offered the following features and materials (screenshots are provided in Multimedia Appendix 2):

- *Study information and participant consent pages*. This page provided the basic information of the study; the user could indicate their consent for participation in the study at the end of the page.
- *Survey of participants' demographic information, living situation, initial motivation level, and exercise habits*. This section allowed users to skip any question they did not want to answer.
- *MBTI survey*. This page provided a questionnaire to determine personality type.
- *Exercise plans and alarms*. Users could enter any number of exercises they wanted to perform every day, such as running for 10 minutes at 6 AM and running for 15 minutes at 6 PM. They could also decide the number of days they wanted to exercise per week. We allowed users to set their goals to ensure that they were tailored to their health conditions. We defined adherence to exercise plans as adherence to users' own plans instead of a universal plan, as, in reality, users have very different health and living conditions, and it is not realistic to establish a universal plan.
- *Display of motivational messages* (settings vary based on the experimental groups, as described in the *Field Experiment* section). An alarm would ring at the time the user set for exercise; a message would be pushed to the main screen (unless the user was in a group that did not receive messages). The alarm could be snoozed 3 times for 5 minutes each. For the groups with repetition, the first notification (with or without a message) appeared 30 minutes in advance to remind users of their plans; then, the reminder appeared again at scheduled times.
- *Record of the exercise plan was achieved*. The app asked whether the exercise plans were followed a few hours after the scheduled time, and users could click *yes* or *no* for each exercise item.

Stage 3: Field Experiment

In this research, we designed a field experiment and received approval from the institutional review board of the University of Illinois. We conducted a field experiment using the mHealth app ActiveTrack. ActiveTrack is an exercise planning mobile app designed based on our theory-based message tailoring method that aims to examine the effects of (1) tailored communication, (2) motivational messages, and (3) repetition of messages on users' adherence to their goal settings. ActiveTrack was available to anyone for download from the Apple store. We also promoted our app through gyms, where many members participated in weight loss programs. Our partnered trainers introduced ActiveTrack to their trainees but were not authorized to monitor their trainees' behavior on the app. We targeted users who used the app for at least two months. The experiment included the following steps:

- Participants downloaded the research app (the user needed to have an iPhone).
- Participants were randomly assigned following simple randomization procedures (computerized random numbers) to 1 of 12 groups when they downloaded the app (randomization was achieved by embedding the random assignment process in the app development). The groups were defined as follows:
  - personality type: thinking type or feeling type
  - message type: no message, emotional message, or logical message
  - repetition: not repeated or repeated once (ie, 1 reminder 30 minutes before the scheduled alarm)

Therefore, we had a 2×3×2 design with 12 groups. Participants remained blinded to the assigned intervention to prevent them from being influenced by such knowledge;

- Participants were presented with the study's information (ie, purpose, procedures, voluntary nature, confidentiality, risks and benefits, and contact information) and signed a consent form (the declaration can be found in Multimedia Appendix 3).
- Participants provided background information such as age, sex, race, weight, and exercise times per week through the app (the information page is available in Multimedia Appendices 2 and 4).
- Participants completed the MBTI questions through the app to allow us to determine their personality types. As we focused on only one dimension of the MBTI test, as described in the *Tailored Communication* section, we selected only the questions that could determine a person's thinking or feeling personalities (MBTI assessment in Multimedia Appendices 2 and 5).
- Participants entered their exercise plans (open-ended text entry) and alarms (specific time). Exercise could be performed any number of times per week and per day based on individual health conditions. Example screenshots can be found in Multimedia Appendix 2.

- Participants received tailored messages. On the basis of our 2×3×2 research design, there were 12 scenarios. Participants with thinking or feeling personality types may receive emotional, logical, or no messages with or without message repetition. Examples of the message display can be found in Multimedia Appendix 2.
- Participants used the app and self-reported whether each exercise item was completed (completion rate). Example screenshots are shown in Multimedia Appendix 2.
- All user inputs, alarms, and reminders were recorded in a web-based database. For privacy protection, we did not collect information that could be used to identify users.

Results

Participants

In total, 160 users downloaded the app during the 2-month period; after excluding users who stopped using the app in the middle of the process, 89 participants remained. The attrition rate was 44.4% (71/160). We compared the characteristics of those who dropped out (inactive users) and those who remained (active users) and did not find any significant differences. Of these active participants, 55 were female and 31 were male (3 did not indicate their sex). Participants' ages ranged from 19 to 56 years, with a mean age of 28.8 (SD 8.3) years. Among all active users, results showed no significant difference in adherence level across sexes and ages. In addition, we found that white-collar workers had 17% higher adherence level to exercise plans than blue-collar workers (t_{63}=−2.045, two-tailed; P=.04). Moreover, higher exercise frequencies per week were associated with a higher adherence level (t_{63}=2.341, two-tailed; P=.02). We further tested the difference in adherence between those who downloaded the app following their trainers' recommendations and those who downloaded the app voluntarily, and the results showed no significant differences between the 2 groups (t_{63}=−0.254, two-tailed; P=.80).

We also compared the characteristics of those who dropped out (inactive users) and those who remained (active users). We found no significant difference in age and exercise frequency between the active and inactive groups. We used a chi-square test to examine any differences between active and inactive users in sex, job type, and working hours. There were no significant differences in sex and working hours between the 2 groups. However, in the active group, we found that white-collar workers had a higher adherence rate than blue-collar workers. In the inactive group, white-collar workers were more likely to quit. The job type was a critical factor that affected users' retaining behavior.

The sample size was relatively small because of the difficulty of recruiting active users; however, the sample size was similar to that in research in the health care domain, which investigated the effectiveness of using mobile phones for health management. In a review of the effectiveness of mHealth and technology-based health behavior management interventions, all 7 studies related to physical activity behaviors included 17-150 participants [72]. Payne et al [13] systematically searched and described the literature on mobile apps used in health behavior interventions; 17 of the 24 studies reviewed had a sample of fewer than 100 participants. In addition, there was a limited impact of the low power caused by the small sample size in this study, as detailed in the following sections. The number of active participants in each group and the mean and SD of plan achievement are presented in Table 1.

Table 1. Descriptive statistics of the active participants in each group.

Personality and message type	Message repetition	Adherence rate, mean (SD)	Active participants, n (%)
Thinking			
None			
	None	41.67 (45.78)	12 (13)
	Once	77.31 (29.72)	14 (16)
Logical			
	None	57.51 (46.10)	7 (8)
	Once	92.99 (14.54)	9 (10)
Emotional			
	None	73.80 (17.18)	11 (12)
	Once	86.58 (16.40)	11 (12)
Feeling			
None			
	None	83.33 (28.87)	3 (3)
	Once	91.67 (16.67)	4 (4)
Logical			
	None	95.24 (8.25)	3 (3)
	Once	83.73 (21.10)	6 (7)
Emotional			
	None	91.75 (10.45)	3 (3)
	Once	94.21 (9.48)	6 (7)

Model Testing Results

Overview

To examine the effects of personality type, message type, and repetition of reminders as well as the two-way interaction effects between the variables on adherence to exercise plans, we used a three-way analysis of variance. To meet the assumption of homogeneity of variance of error, we used a square operation to transform the dependent variable. The *F* test for heteroskedasticity suggests that the equal variance of the error assumption is met ($F_{1,85}$=1.26; *P*=.26). The results of the analysis of variance model are presented in Table 2. The results suggested a good overall fit of the main effects of personality type and repetition and a significant interaction between personality type and repetition. To understand the main effects, we conducted a Tukey honest significant difference posthoc analysis to compare the mean difference across the different groups; the results are presented in Table 2 and Figures 2-4.

Table 2. Model testing results.[a]

Source	Type III sum of squares	Mean square	*F* test (*df*)	*P* value
Personality type	73,471,244	73,471,244	6.537 (1)	.01
Message type	32,216,925	16,108,463	1.433 (2)	.24
Repetition	93,902,535	93,902,535	8.355 (1)	.005
Personality type×message type	13,568,399	6,784,200	0.604 (2)	.55
Personality type×repetition	38,011,469	38,011,469	3.382 (1)	.07
Message type×repetition	3,427,468	1,713,734	0.152 (2)	.86
Personality type×message type×repetition	12,183,242	6,091,621	0.542 (2)	.85
Residuals	865,416,174	11,239,171	N/A[b]	N/A

[a]R^2=0.28; adjusted R^2=0.20.

[b]N/A: not applicable.

Figure 2. Personality type main effect.

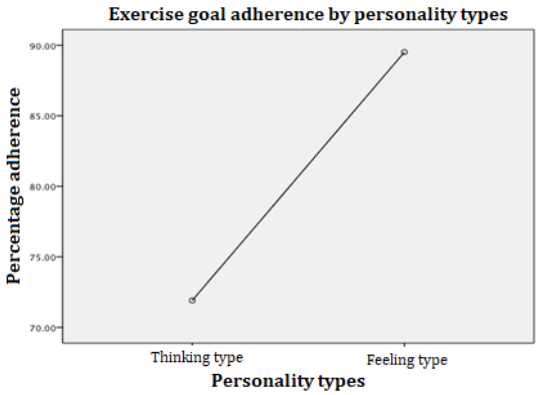


Figure 3. Repetition main effect.

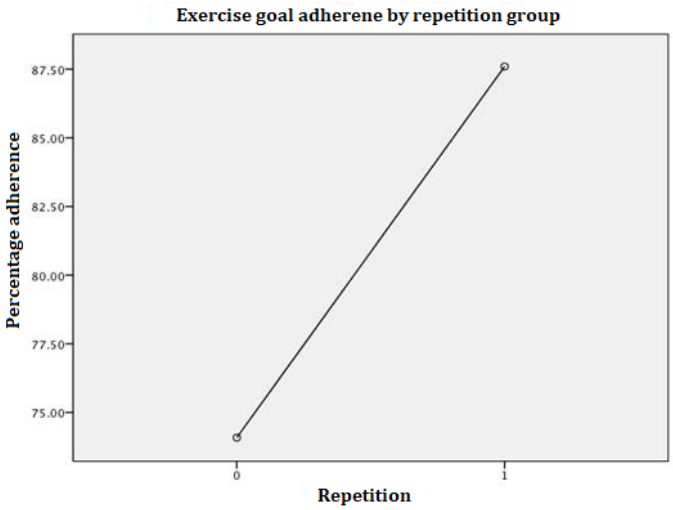
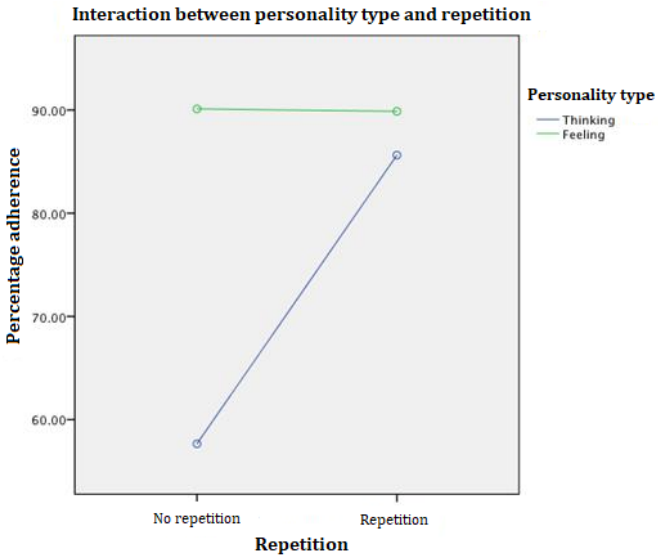


Figure 4. Interaction effect of personality type and repetition.



Personality Main Effect

From Figure 2, we can see that independent of other variables, people with a feeling personality type were significantly more likely to adhere to exercise plans. This is an interesting finding that suggests that the effectiveness of mobile apps on influencing behavior might be different for people with different personality types. Two-sample *t* test of mean difference of the dependent variable by personality types shows that the mean of the group with thinking type personality (mean 71.65%, SD 34.24%) is significantly lower (95% CI −31.28 to −5.03; *P*=.007) than that of the feeling type personality group (mean 89.79%, SD 16.39%).

Repetition Main Effect

Figure 3 shows that participants who received repeated messages had significantly increased adherence to exercise plans. Two-Sample *t* test of mean difference of the dependent variable by repetition shows that the mean of the group with one more notification (mean 86.12%, SD 20.92%) is significantly higher (95% CI 7.78-31.56; *P*=.001) than that of the group without repetition (mean 63.38%, SD 38.33%).

Interaction Between Personality Type and Repetition

We then investigated the significant interaction between personality type and repetition, as reported in Table 3 and Figure 4. The results suggested that (1) when there is no repetition, people with a feeling personality type have a significantly higher completion rate than people with a thinking personality type and, (2) when there is repetition, the completion rate of people with a thinking personality type significantly improves, there is no significant change in the completion rate of people with a feeling personality type, and the 2 groups have no significant difference in terms of completion rate. This is an interesting finding; combined with a previous finding, it suggests that although people with a thinking personality type might have less commitment to the plans they set for themselves, with repeated reminders and messages, these people can achieve a similar completion rate as those with a feeling personality type. This might be because the *mere exposure* effect impacts a person with a thinking personality type more than a person with a feeling personality type.

Table 3. Comparison of personality type and repetition groups based on estimated marginal means.

Group compared	Mean difference (%)	*P* value	95% CI
Thinking, repetition-thinking, no repetition	27.34	.001	8.95 to 45.73
Feeling, repetition-feeling, no repetition	−0.83	.99	−31.41 to 29.75
Feeling, no repetition-thinking, no repetition	33.21	.01	5.31 to 61.1
Feeling, repetition-thinking, repetition	5.03	.93	−17.21 to 27.28
Feeling, repetition-thinking, no repetition	32.38	<.001	9.65 to 55.10
Thinking, repetition-feeling, no repetition	−5.86	.94	−33.38 to 21.64

Summary of Results

Our results suggest that mHealth apps can be effective in promoting adherence. People with a feeling personality type were more likely to adhere to exercise plans than those with a thinking personality type. When receiving repeated reminders and messages, the adherence rates of people with a thinking personality type significantly improved and showed no significant difference from those of people with a feeling personality type. These results are consistent with prior studies that suggest that different personality traits are linked to certain behavioral tendencies, such as proneness to addiction [73] and excessive use of the internet [74]. Our study suggests that personality can contribute to the persistence of adherence to exercise plans.

Additional Findings

Although the purpose of this research was to examine the effects of tailored messages based on personality and repetition, other tailoring criteria might be selected to motivate exercise behaviors. For example, demographic tailoring has been found to have effects that are independent of theory-based tailoring [75,76]. To further enhance our understanding of the effects of message tailoring and justify the applicability of our model, we conducted an additional analysis to investigate the interactions between sex and message type. We found that emotional messages to women and logical messages to men had a 35.56% higher adherence level (t_{46}=3.278; *P*=.002) than the other way around (emotional messages to men and logical messages to women). These findings warrant further analysis.

Robustness Check

In the behavior research field, the issue of violation of parametric tests' statistical assumptions is rather common, such as skewed distribution, heteroskedasticity, and violation of the independence of errors. Consequently, nonparametric tests have been proposed and widely used to address these issues because of their advantage of not being limited to the assumptions of distribution or homogeneity and because they can be applied to a small sample size [77]. Resampling methods, permutation (randomization test), and bootstrapping are common nonparametric methods.

Permutation tests use all possible permutations of a treatment variable or dependent variable, whereas all other independent variables are fixed to construct the exact null distribution using the available data to determine how extreme the observed test statistic of research interest is against the null distribution. The randomization test [78] relies on the same idea as permutation tests; however, these tests compare the observed statistic against the approximate null distribution generated by repeating a large number of permutations (eg, 10,000-time default in the R package) rather than all possible permutations. As the null distribution is generated empirically from the observed samples and makes no assumptions regarding the population, permutation (randomization test) is beneficial for evaluating the statistical significance or treatment effect of any variable of interest. Another resampling method, which is bootstrap, repeats the available sample and draws from the same sample with replacement to calculate the test statistics and construct an empirical distribution. In experiments, the bootstrap method is a useful way to examine the treatment effects of a designed experiment [79,80].

We used a randomization test and bootstrap resampling method to assess the main effects of the 3 factors and their interactions. By permuting 10,000 times in the randomization test and resampling with replacement 10,000 times for the bootstrap test, our robustness check of the treatment effects shows that the 2 main effects of personality and repetition and the interaction effect between personality and repetition were significant. The results are summarized in Table 4.

Table 4. Comparison of resampling methods.

Type of treatment	Base model[a], *P* value	Randomization test, *P* value	95% bootstrap CI
Emotional message	.24	.14	−17.63 to 55.49
No message	.24	.16	−56.63 to 30.22
Personality	.06	.03	0.07 to 77.89
Repetition	.02	.01	0.85 to 74.48
Emotional message×personality	.65	.30	−83.03 to 17.73
No message×personality	.88	.45	−61.64 to 46.59
Emotional message×repetition	.23	.13	−65.14 to 13.05
No message×repetition	.99	.51	−49.51 to 45.07
Personality×repetition	.06	.06	−91.90 to −6.79
Emotional message×personality×repetition	.42	N/A[b]	−9.81 to 81.93
No message×personality×repetition	.57	N/A	−36.88 to 85.84

[a]Base message type in the regression model is a logical message.

[b]N/A: not applicable.

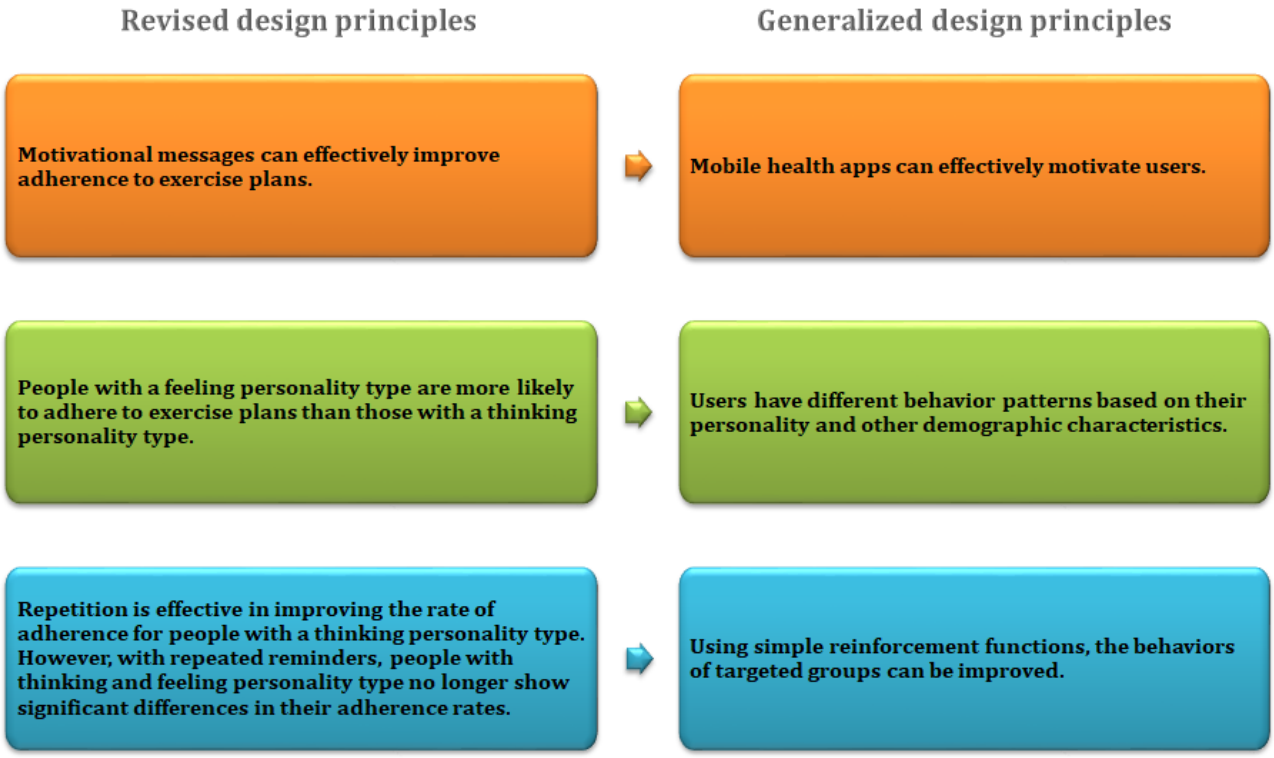
We also used a randomization test and bootstrap resampling method to examine the effect of *sex-message tailoring*. The effect was significant, with a 95% CI between 17.07% and 55.74% and *P*=.004 in the randomization test. Both resampling methods show consistent results in terms of the treatment effect.

Discussion

Learning With ADR

According to our research process indicated in Figure 1, the evidence collected from the field experiment allowed us to summarize the insights presented in Figure 5. The design principles were revised based on the knowledge from the design and experimental effort and are described in Figure 5. These design principles were used to generate a series of app development problems, such as motivation for other behavioral changes and apps adapted to individual differences. These generalized principles are also presented in Figure 5.

Figure 5. Revised design principles and generalizations.



Research Implications

In this research, we enhanced the ADR framework by developing the ActiveTrack iOS app and incorporating a behavior experiment into the process. When participants downloaded the app, they were randomly assigned to one of the groups for the behavioral experiment. On the basis of the users' real-time responses, we revised the functions and designs of the app. As a result of the field experiment, we updated the design principles through the ADR process. Our field experiment results suggested that (1) the mobile app was effective in motivating adherence to exercise plans among users with both thinking and feeling personality types, (2) personality type was associated with the likelihood of adherence to exercise plans, and (3) repetition improved the exercise plan adherence rate of people with a thinking-type personality, and with repeated reminders, the adherence rate of people with a thinking personality type becomes similar to that of people with a feeling personality type. Our additional findings further imply that users' behavior may depend on their (1) job types, (2) exercise habits, and (3) gender. By adding the behavior experiment as a component of ADR, we were able to learn design principles from the behavior experiment results and generalize these principles to a class of design problems: how to effectively motivate users through the personalized design of apps.

This study makes several contributions to the field. First, this research contributes to the literature on mHealth by investigating the effectiveness of mobile apps in motivating health behaviors. Second, this research improves our understanding of the impact of individual personality, message type, and repetition on the effectiveness of motivation. Specifically, through the field experiment, we were able to follow the participants and capture actual behavior change over time, which has been a challenge in behavior-related studies [81]. Third, field surveys, behavior intervention experiments, and the design science approach were normally used separately in the research. In this study, we used all of the above approaches to address the research questions. Combining these methods enabled us to actively design the app by observing usage data and using the design artifact to carry out the research. Thus, by incorporating a field experiment into the research process, we propose an enhanced ADR process. This research process enables us to better connect research with practical and relevant problems. Finally, the study provides several design principles that can be applied to the design of other mHealth and motivational apps, such as apps for lifestyle management, energy-saving behaviors, and adherence-to-treatment plans.

Lessons Learned From the Apple ResearchKit

In this research, we used the Apple ResearchKit to build artifacts. Apple released ResearchKit, an open-source software framework for medical research, in March 2015. The ResearchKit framework comes with some predefined modules commonly used in research procedures, such as informed consent, surveys, and active tasks, to make the development of mobile apps feasible and more convenient for research. This framework has gained much attention from research institutions and companies owing to its global release [82-84]. However, using ResearchKit and similar development frameworks is still a novel method for scientific research projects, and there are many challenges in conducting research using this method. Through this project, we learned that some of those challenges include difficulty promoting the app because of the number of apps available in the same category in the Apple store, verifying the validity of inputs, and ensuring data security and privacy. In future research, some of these challenges can be overcome

during the study design phase. For example, using the heart rate module, the app could more accurately estimate exercise completion. If a study targets a specific group of users, such as people with certain types of disease, it would be a good practice to promote the use of the app among targeted audiences, such as online patient support groups. It is convenient to download the app from the app store, so the barrier for a targeted population to use the app is low regardless of their physical locations.

Limitations and Future Research

This study has several limitations that can be addressed and improved in future studies. First, the sample size of this research was relatively small because of the difficulty in recruiting active users over the 2 months. However, the sample size was in line with that of a similar study. To address this issue, we used both a randomization test and bootstrap resampling methods to test our treatment effects. The consistent results showed that the main effects of personality and repetition and the two-way interaction between personality and repetition were statistically significant. Future researchers can collaborate with developers, health care providers, and health social networks to increase the potential of participant recruitment. Second, self-reported adherence might be biased. Although we implemented strategies such as reminding participants that the data would not be shared with anyone and honest responses would improve the design of the messages, the validity of self-reported exercise levels could not be verified. Some smartphone features, such as the tracking of heartbeats, can be used for more accurate estimates of exercise in future studies. Third, we attempted to limit the impacts of other confounding factors that may influence the behavior of the users through random assignment to groups; however, there might still have been influences from confounding factors. This is a difficulty that similar research has always faced, as it is difficult to make behavior change in such experiments totally independent of external factors [81].

Potential topics for future research include other strategies for personalized app design and motivation, such as using different theory-based tailoring methods, testing strategies to retain users by analyzing their churn behavior, and identifying ways to use the data collected from these apps to promote health and well-being (ie, different metrics to measure users' lifestyle change). For example, we found that 14% (12/84) of the users achieved higher average exercise frequencies than their original exercise habits after 2 months. This finding implies that users may change their lifestyles using our mHealth app. Future research can also focus on the clinical validation of the experimental results, ways of using economic incentives to promote healthy lifestyles, new business models for health care providers and insurance companies to motivate adherence, and the use of other wearable devices, such as smartwatches, to motivate health behaviors.

Acknowledgments

The authors would like to acknowledge the financial support from the State Farm Foundation.

Conflicts of Interest

RTS is the developer of the research app ActiveTrack. All other authors have no conflicts to declare.

Multimedia Appendix 1
Emotional and logical messages.
[DOCX File , 44 KB - mhealth_v9i6e19941_app1.docx]

Multimedia Appendix 2
Screenshots of the design artifact—ActiveTrack.
[PNG File , 149 KB - mhealth_v9i6e19941_app2.png]

Multimedia Appendix 3
Consent form declaration.
[PDF File (Adobe PDF File), 79 KB - mhealth_v9i6e19941_app3.pdf]

Multimedia Appendix 4
Questionnaire for demographics information.
[PDF File (Adobe PDF File), 67 KB - mhealth_v9i6e19941_app4.pdf]

Multimedia Appendix 5
Myers-Briggs Type Indicator assessment.
[PDF File (Adobe PDF File), 87 KB - mhealth_v9i6e19941_app5.pdf]

References

1. Physical activity. U.S. Department of Health and Human Services. URL: https://www.healthypeople.gov/2020/topics-objectives/topic/physical-activity [accessed 2020-04-20]
2. Booth F, Roberts C, Laye M. Lack of exercise is a major cause of chronic diseases. Compr Physiol 2012 Apr;2(2):1143-1211 [FREE Full text] [doi: 10.1002/cphy.c110025] [Medline: 23798298]
3. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. J Am Med Assoc 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
4. Park A. Lack of exercise as deadly as smoking, study finds. Time. 2012. URL: https://healthland.time.com/2012/07/18/lack-of-exercise-as-deadly-as-smoking-study-finds/ [accessed 2020-04-20]
5. Coulson J, McKenna J, Field M. Exercising at work and self‐reported work performance. Intl J of Workplace Health Mgt 2008 Sep 26;1(3):176-197. [doi: 10.1108/17538350810926534]
6. Patient adoption of mHealth. IMS Institute for Healthcare Informatics. 2015. URL: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf [accessed 2020-04-20]
7. The growing value of digital health. The IQVIA Institute Reports. 2017. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health [accessed 2020-04-20]
8. Healthcare consumers in a digital transition. Rock Health. 2021. URL: https://rockhealth.com/reports/healthcare-consumers-in-a-digital-transition/ [accessed 2020-04-20]
9. Maddison R, Rawstorn JC, Islam SM, Ball K, Tighe S, Gant N, et al. mHealth interventions for exercise and risk factor modification in cardiovascular disease. Exerc Sport Sci Rev 2019 Apr;47(2):86-90 [FREE Full text] [doi: 10.1249/JES.0000000000000185] [Medline: 30883472]
10. Oinas-Kukkonen H, Harjumaa M. Towards deeper understanding of persuasion in software and information systems. In: Proceedings of the First International Conference on Advanced in Computer-Human Interaction. 2008 Presented at: First International Conference on Advances in Computer-Human Interaction; February 10-15, 2008; Sainte Luce, Martinique, France p. 200-205. [doi: 10.1109/achi.2008.31]
11. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013 Feb;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
12. Marcano BJ, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013;11:CD010013. [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
13. 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;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
14. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. PLoS One 2016 Nov;11(11):e0166718 [FREE Full text] [doi: 10.1371/journal.pone.0166718] [Medline: 27861583]
15. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24:486-500. [doi: 10.17705/1CAIS.02428]
16. Geuens J, Swinnen TW, Westhovens R, de VK, Geurts L, Vanden AV. A review of persuasive principles in mobile apps for chronic arthritis patients: opportunities for improvement. JMIR Mhealth Uhealth 2016 Oct 13;4(4):e118 [FREE Full text] [doi: 10.2196/mhealth.6286] [Medline: 27742604]
17. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
18. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. J Med Syst 2016 Mar;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
19. Maibach EW, Weber D, Massett H, Hancock GR, Price S. Understanding consumers' health information preferences: development and validation of a brief screening instrument. J Health Commun 2006 Dec;11(8):717-736. [doi: 10.1080/10810730600934633] [Medline: 17190779]
20. Peppers D, Rogers M, Dorf B. Is your company ready for one-to-one marketing? Harv Bus Rev 1999;77(1):151-160. [Medline: 10345390]
21. Storey J, Saffitz G, Rimon J. Social marketing. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research and Practice. 4th Ed. Hoboken, New Jersey, United States: John Wiley & Sons; 2008:435-464.
22. Strecher V, Rimer B, Monaco K. Development of a new self-help guide--Freedom From Smoking for you and your family. Health Educ Q 1989 Feb;16(1):101-112 [FREE Full text] [doi: 10.1177/109019818901600111] [Medline: 2703347]
23. Robinson RG, Sutton CD, James DA, Orleans CT. Pathways to freedom: Winning the fight against tobacco. US Department of Health and Human Services, Centers for Disease Control and Prevention. URL: https://www.cdc.gov/tobacco/quit_smoking/how_to_quit/pathways/pdfs/pathways.pdf [accessed 2020-04-20]
24. Rimer B, Orleans C, Fleisher L, Cristinzio S, Resch N, Telepchak J, et al. Does tailoring matter? The impact of a tailored guide on ratings and short-term smoking-related outcomes for older smokers. Health Educ Res 1994 Mar;9(1):69-84. [doi: 10.1093/her/9.1.69] [Medline: 10146734]

25. Davis SW, Cummings KM, Rimer BK, Sciandra R, Stone JC. The impact of tailored self-help smoking cessation guides on young mothers. Health Educ Q 1992 Sep 04;19(4):495-504. [doi: 10.1177/109019819201900412] [Medline: 1452449]
26. Keintz MK, Fleisher L, Rimer BK. Reaching mothers of preschool-aged children with a targeted quit smoking intervention. J Community Health 1994 Feb;19(1):25-40. [doi: 10.1007/bf02260519]
27. Jung CG. Psychiatric studies. In: Read H, Fordham M, Adler G, Hull RF, editors. The Collected Works of CG Jung, Vol. 1. 2nd Edition. Princeton, New Jersey, United States: Princeton University Press; 1992:1-288.
28. Myers IB, McCaulley MH, Quenk NL, Hammer AL. MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator, 3rd Edition. Minnesota, United States: Consulting Psychologists Press; 1998:1-420.
29. Graff WS. The effectiveness of systematic desensitization in the reduction of test anxiety in Jungian thinking versus feeling personality types. ProQuest Information & Learning. United States -- Alabama: Auburn University; 1976. URL: https://tinyurl.com/yuyjrnv3 [accessed 2021-04-30]
30. Carskadon TG. Clinical and counseling aspects of the Myers-Briggs Type Indicator: a research review. Research in Psychological Type. 1979. URL: https://www.researchgate.net/publication/286364037_Clinical_and_counseling_aspects_of_the_Myers-Briggs_Type_Indicator_A_research_review [accessed 2021-05-04]
31. Jinkerson J, Masilla A, Hawkins RC. Can MBTI dimensions predict therapy outcome: differences in the thinking-feeling function pair in CBT. Res Psychothe Psychopathol Proc Outcome 2015 Mar 11;18(1):A. [doi: 10.4081/ripppo.2015.167]
32. Arain AA. Relationships among counseling clients' personalities, expectations,problems (Doctoral dissertation). Rutgers University, New Brunswick, NJ. 1967. URL: https://search.proquest.com/openview/e286ee4d35edb5581a407b7f26fdd554/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2021-05-04]
33. Garden A. Relationships between MBTI profiles, motivation profiles, and career paths. J Psychol Type. 1997. URL: https://www.capt.org/jpt/pdfFiles/Garden_A_Vol_41_3_16.pdf [accessed 2021-05-04]
34. DeBono KG, Harnish RJ. Source expertise, source attractiveness, and the processing of persuasive information: a functional approach. J Pers Soc Psychol 1988 Oct;55(4):541-546. [doi: 10.1037/0022-3514.55.4.541]
35. Klein WM, Harris PR. Self-affirmation enhances attentional bias toward threatening components of a persuasive message. Psychol Sci 2009 Dec;20(12):1463-1467. [doi: 10.1111/j.1467-9280.2009.02467.x] [Medline: 19883488]
36. Singh SN, Cole CA. The effects of length, content, and repetition on television commercial effectiveness. J Mark Res 2018 Dec 20;30(1):91-104. [doi: 10.1177/002224379303000108]
37. Malaviya P. The moderating influence of advertising context on ad repetition effects: the role of amount and type of elaboration. J Consum Res 2007 Jun 01;34(1):32-40. [doi: 10.1086/513044]
38. Lewis IM, Watson B, White KM, Tay R. Promoting public health messages: should we move beyond fear-evoking appeals in road safety? Qual Health Res 2007 Jan;17(1):61-74. [doi: 10.1177/1049732306296395] [Medline: 17170244]
39. Gifford R, Comeau LA. Message framing influences perceived climate change competence, engagement, and behavioral intentions. Glob Environ Change 2011 Oct;21(4):1301-1307. [doi: 10.1016/j.gloenvcha.2011.06.004]
40. Prentice-Dunn S, Floyd D, Flournoy JM. Effects of persuasive message order on coping with breast cancer information. Health Educ Res 2001 Mar;16(1):81-84. [doi: 10.1093/her/16.1.81] [Medline: 11252286]
41. Sherman DK, Mann T, Updegraff JA. Approach/avoidance motivation, message framing, and health behavior: understanding the congruency effect. Motiv Emot 2006 Jun 2;30(2):165-169 [FREE Full text] [doi: 10.1007/s11031-006-9001-5] [Medline: 19079797]
42. Prestwich A, Perugini M, Hurling R. Can the effects of implementation intentions on exercise be enhanced using text messages? Psychol Health 2009 Jul;24(6):677-687. [doi: 10.1080/08870440802040715] [Medline: 20205020]
43. Quilici J, Fugon L, Beguin S, Morange PE, Bonnet J, Alessi M, et al. Effect of motivational mobile phone short message service on aspirin adherence after coronary stenting for acute coronary syndrome. Int J Cardiol 2013 Sep 20;168(1):568-569. [doi: 10.1016/j.ijcard.2013.01.252] [Medline: 23462636]
44. Pallak SR, Murroni E, Koch J. Communicator attractiveness and expertise, emotional versus rational appeals, and persuasion: a heuristic versus systematic processing interpretation. Soc Cogn 1983 Jun;2(2):122-141. [doi: 10.1521/soco.1983.2.2.122]
45. Miceli M, Rosis FD, Poggi I. Emotional and non-emotional persuasion. Appl Artif Intell 2006 Jun;20(10):849-879. [doi: 10.1080/08839510600938193]
46. Tykocinskl O, Higgins ET, Chaiken S. Message framing, self-discrepancies, and yielding to persuasive messages: the motivational significance of psychological situations. Pers Soc Psychol Bull 2016 Jul 02;20(1):107-115. [doi: 10.1177/0146167294201011]
47. Shen L, Dillard JP. Message frames interact with motivational systems to determine depth of message processing. Health Commun 2009 Sep 31;24(6):504-514. [doi: 10.1080/10410230903104897] [Medline: 19735028]
48. Toulmin SE. The Uses of Argument. Cambridge, United Kingdom: Cambridge University Press; 2003.
49. Kim D, Benbasat I. The effects of trust-assuring arguments on consumer trust in internet stores: application of toulmin's model of argumentation. Inf Syst Res 2006 Sep;17(3):286-300. [doi: 10.1287/isre.1060.0093]
50. Clore GL, Gasper K. Feeling is believing: Some affective influences on belief. In: Emotions and Beliefs. Cambridge, United Kingdom: Cambridge University Press; 2000:10-44.

51. Murphy ST, Zajonc RB. Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. J Pers Soc Psychol 1993;64(5):723-729. [doi: 10.1037/0022-3514.64.5.723]
52. Cacioppo JT, Petty RE. Effects of message repetition and position on cognitive response, recall, and persuasion. J Pers Soc Psychol 1979 Jan;37(1):97-109. [doi: 10.1037/0022-3514.37.1.97]
53. Burgoon M, Miller MD. Overcoming resistance to persuasion VIA contiguous reinforcement and repetition of message. Psychol Rep 2016 Aug 31;66(3):1011-1022. [doi: 10.2466/pr0.1990.66.3.1011]
54. Zajonc R. Mere exposure: a gateway to the subliminal. Curr Dir Psychol Sci 2016 Jun 23;10(6):224-228. [doi: 10.1111/1467-8721.00154]
55. Saegert SC, Jellison JM. Effects of initial level of response competition and frequency of exposure on liking and exploratory behavior. J Pers Soc Psychol 1970 Nov;16(3):553-558. [doi: 10.1037/h0029952] [Medline: 5497259]
56. Johnson HH, Watkins TA. The effects of message repetitions on immediate and delayed attitude change. Psychon Sci 2013 Nov 5;22(2):101-103. [doi: 10.3758/bf03332515]
57. Weiss RF. Role playing and repetition effects on opinion strength. J Soc Psychol 1971 Oct;85(1):29-35. [doi: 10.1080/00224545.1971.9918541] [Medline: 5114721]
58. McCullough JL, Ostrom TM. Repetition of highly similar messages and attitude change. J Appl Psychol 1974;59(3):395-397. [doi: 10.1037/h0036658]
59. Smith GF, Dorfman DD. The effect of stimulus uncertainty on the relationship between frequency of exposure and liking. J Pers Soc Psychol 1975;31(1):150-155. [doi: 10.1037/h0076174]
60. Harrison A. Mere exposure. Adv Exp Soc Psychol 1977;10:39-83. [doi: 10.1016/S0065-2601(08)60354-8]
61. Bettinghaus EP. Health promotion and the knowledge-attitude-behavior continuum. Prev Med 1986 Sep;15(5):475-491. [doi: 10.1016/0091-7435(86)90025-3]
62. Brooks ME, Highhouse S. Familiarity breeds ambivalence. Corp Reputation Rev 2006 Aug 31;9(2):105-113. [doi: 10.1057/palgrave.crr.1550016]
63. Fang X, Singh S, Ahluwalia R. An examination of different explanations for the mere exposure effect. J Consum Res 2007 Jun 01;34(1):97-103. [doi: 10.1086/513050]
64. Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educ Psychol Meas 2016 Jul 02;60(6):821-836. [doi: 10.1177/00131640021970934]
65. Pohl RF. Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory, First Edition. London: Psychology Press; 2005:1-450.
66. Serenko A, Bontis N. What's familiar is excellent: The impact of exposure effect on perceived journal quality. J Informetr 2011 Jan;5(1):219-223. [doi: 10.1016/j.joi.2010.07.005]
67. Malone TW. How do people organize their desks? ACM Trans Inf Syst 1983 Jan;1(1):99-112. [doi: 10.1145/357423.357430]
68. Lord S, Ward J, Williams P, Strudwick M. The effect of a 12-month exercise trial on balance, strength, and falls in older women: a randomized controlled trial. J Am Geriatr Soc 1995 Nov;43(11):1198-1206. [doi: 10.1111/j.1532-5415.1995.tb07394.x] [Medline: 7594152]
69. Estabrooks P, Carron A. Group cohesion in older adult exercisers: prediction and intervention effects. J Behav Med 1999;22(6):575-588. [doi: 10.1023/A:1018741712755]
70. Sein MK, Henfridsson O, Purao S, Rossi M, Lindgren R. Action design research. MIS Q 2011;35(1):37. [doi: 10.2307/23043488]
71. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
72. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013 Jan;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
73. Müller SE, Weijers H, Böning J, Wiesbeck GA. Personality traits predict treatment outcome in alcohol-dependent patients. Neuropsychobiology 2008 Jul 24;57(4):159-164. [doi: 10.1159/000147469] [Medline: 18654085]
74. Munno D, Cappellin F, Saroldi M, Bechon E, Guglielmucci F, Passera R, et al. Internet addiction disorder: personality characteristics and risk of pathological overuse in adolescents. Psychiatry Res 2017 Mar;248:1-5. [doi: 10.1016/j.psychres.2016.11.008] [Medline: 27988425]
75. Noar SM, Harrington NG, Aldrich RS. The role of message tailoring in the development of persuasive health communication messages. Annal Int Commun Asso 2016 May 18;33(1):73-133. [doi: 10.1080/23808985.2009.11679085]
76. Harrington NG, Noar SM. Reporting standards for studies of tailored interventions. Health Educ Res 2012 Apr 08;27(2):331-342. [doi: 10.1093/her/cyr108] [Medline: 22156230]
77. Peres-Neto PR, Olden JD. Assessing the robustness of randomization tests: examples from behavioural studies. Anim Behav 2001 Jan;61(1):79-86. [doi: 10.1006/anbe.2000.1576] [Medline: 11170698]
78. Edgington ES. Randomization tests. In: Lovric M, editor. International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer; 2014:A.
79. Davison AC, Hinkley DV. Bootstrap Methods and Their Application (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press; 1997:1-594.

80. Kenett RS, Rahav E, Steinberg DM. Bootstrap analysis of designed experiments. Qual Reliab Engg Int 2006 Oct;22(6):659-667. [doi: 10.1002/qre.802]
81. Andrews PY. System personality and persuasion in human-computer dialogue. ACM Trans Interact Intel Syst 2012 Jun;2(2):1-27. [doi: 10.1145/2209310.2209315]
82. Eisenstein M. GSK collaborates with Apple on ResearchKit. Nat Biotechnol 2015 Oct 8;33(10):1013-1014. [doi: 10.1038/nbt1015-1013a] [Medline: 26448070]
83. Webster DE, Suver C, Doerr M, Mounts E, Domenico L, Petrie T, et al. The Mole Mapper Study, mobile phone skin imaging and melanoma risk data collected using ResearchKit. Sci Data 2017 Feb 14;4(1):170005 [FREE Full text] [doi: 10.1038/sdata.2017.5] [Medline: 28195576]
84. Zens M, Woias P, Suedkamp NP, Niemeyer P. "Back on Track": a mobile App observational study using Apple's ResearchKit framework. JMIR Mhealth Uhealth 2017 Mar 28;5(2):e23 [FREE Full text] [doi: 10.2196/mhealth.6259] [Medline: 28246069]

Abbreviations

ADR: action design research
MBTI: Myers-Briggs Type Indicator
mHealth: mobile health

Edited by G Eysenbach; submitted 06.05.20; peer-reviewed by KT Win, B Oh, P Cook; comments to author 26.10.20; revised version received 12.01.21; accepted 04.04.21; published 02.06.21.

Please cite as:
Sun RT, Han W, Chang HL, Shaw MJ
Motivating Adherence to Exercise Plans Through a Personalized Mobile Health App: Enhanced Action Design Research Approach
JMIR Mhealth Uhealth 2021;9(6):e19941
URL: https://mhealth.jmir.org/2021/6/e19941
doi:10.2196/19941
PMID:34076580

Original Paper

The SleepFit Tablet Application for Home-Based Clinical Data Collection in Parkinson Disease: User-Centric Development and Usability Study

Alessandro Mascheroni[1], MSc; Eun Kyoung Choe[2*], PhD; Yuhan Luo[2*], MSc; Michele Marazza[3*], MSc; Clara Ferlito[4*], MD; Serena Caverzasio[4*], MSc; Francesco Mezzanotte[4]; Alain Kaelin-Lang[4,5,6], MD, PhD; Francesca Faraci[1*], MSc, PhD; Alessandro Puiatti[1*], MSc; Pietro Luca Ratti[4*], MD, PhD

[1]Institute of Information Systems and Networking, Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, Lugano, Switzerland

[2]College of Information Studies, University of Maryland, College Park, MD, United States

[3]Information & Communication Technology, Ente Ospedaliero Cantonale, Bellinzona, Switzerland

[4]Neurocenter of Southern Switzerland, Lugano, Switzerland

[5]Faculty of Biomedical Sciences, University of Southern Switzerland, Lugano, Switzerland

[6]Medical School, University of Bern, Bern, Switzerland

[*]these authors contributed equally

Corresponding Author:
Pietro Luca Ratti, MD, PhD
Neurocenter of Southern Switzerland
Regional Hospital of Lugano, EOC, via Tesserete 46
Lugano, CH-6903
Switzerland
Phone: 41 353 412 71 91
Email: pietroluca.ratti@gmail.com

Abstract

Background: Parkinson disease (PD) is a common, multifaceted neurodegenerative disorder profoundly impacting patients' autonomy and quality of life. Assessment in real-life conditions of subjective symptoms and objective metrics of mobility and nonmotor symptoms such as sleep disturbance is strongly advocated. This information would critically guide the adaptation of antiparkinsonian medications and nonpharmacological interventions. Moreover, since the spread of the COVID-19 pandemic, health care practices are being reshaped toward a more home-based care. New technologies could play a pivotal role in this new approach to clinical care. Nevertheless, devices and information technology tools might be unhandy for PD patients, thus dramatically limiting their widespread employment.

Objective: The goals of the research were development and usability evaluation of an application, SleepFit, for ecological momentary assessment of objective and subjective clinical metrics at PD patients' homes, and as a remote tool for researchers to monitor patients and integrate and manage data.

Methods: An iterative and user-centric strategy was employed for the development of SleepFit. The core structure of SleepFit consists of (1) an electronic finger-tapping test; (2) motor, sleepiness, and emotional subjective scales; and (3) a sleep diary. Applicable design, ergonomic, and navigation principles have been applied while tailoring the application to the specific patient population. Three progressively enhanced versions of the application (alpha, v1.0, v2.0) were tested by a total of 56 patients with PD who were asked to perform multiple home assessments 4 times per day for 2 weeks. Patient compliance was calculated as the proportion of completed tasks out of the total number of expected tasks. Satisfaction on the latest version (v2.0) was evaluated as potential willingness to use SleepFit again after the end of the study.

Results: From alpha to v1.0, SleepFit was improved in graphics, ergonomics, and navigation, with automated flows guiding the patients in performing tasks throughout the 24 hours, and real-time data collection and consultation were made possible thanks to a remote web portal. In v2.0, the kiosk-mode feature restricts the use of the tablet to the SleepFit application only, thus preventing users from accidentally exiting the application. A total of 52 (4 dropouts) patients were included in the analyses. Overall compliance

(all versions) was 88.89% (5707/6420). SleepFit was progressively enhanced and compliance increased from 87.86% (2070/2356) to 89.92% (2899/3224; *P*=.04). Among the patients who used v2.0, 96% (25/26) declared they would use SleepFit again.

Conclusions: SleepFit can be considered a state-of-the-art home-based system that increases compliance in PD patients, ensures high-quality data collection, and works as a handy tool for remote monitoring and data management in clinical research. Thanks to its user-friendliness and modular structure, it could be employed in other clinical studies with minimum adaptation efforts.

Trial Registration: ClinicalTrials.gov NCT02723396; https://clinicaltrials.gov/ct2/show/NCT02723396

(JMIR Mhealth Uhealth 2021;9(6):e16304) doi:10.2196/16304

KEYWORDS

Parkinson disease; ecological momentary assessment; finger-tapping test; subjective scales; sleep diaries; tablet application; home-based system

Introduction

Parkinson disease (PD) is a neurodegenerative disorder affecting 1.5% of subjects aged older than 60 years [1]. Its prevalence in the general population is estimated between 1/10,000 and 4/1000 [2]. A progressively impaired motor function leads to loss of autonomy in daily living and reduced quality of life of PD patients [3]. Besides motor symptoms, PD also features nonmotor symptoms involving, among others, cognition, emotional state, autonomic functions, and sleep. From 65% to 95% of PD patients report disturbed sleep or daytime sleepiness [4,5], which further impairs their quality of life or the quality of life of their families [6,7]. Moreover, a substantial proportion of patients with PD report prominent spontaneous, transitory improvements in mobility after sleep and before taking the first morning dose of dopaminergic medications. This phenomenon is referred to as “sleep benefit” [8].

To characterize the wide variations of motor and nonmotor symptoms within the same day and their day-to-day variability in individual patients, multiple repeated assessments are necessary. Moreover, prospective assessment of subjective symptoms is essential for the collection of data that is unbiased by patients' recall. This issue can be critical in patients with PD, who may have subclinical cognitive dysfunction [9]. In fact, a sizeable proportion of patients tend to over or underestimate their symptom severity in retrospect at hospital consultations [10]. Tracking subjective symptoms and their variation is crucial for clinical follow-up of patients with PD. Information derived from patients’ subjective perceptions is as important as objective motor features to optimally adapt pharmacological and nonpharmacological treatments [11].

In addition, remote monitoring has very recently become a strong and urgent need for clinical care since many national health authorities have encouraged patients with chronic diseases not to attend their follow-up consultations unless seeking urgent care in the attempt to limit the spread of the COVID-19 pandemic.

Ecological momentary assessment is a technique involving repeated, prospective sampling of subjects’ current behaviors or experiences in real time in his or her natural environment. Ecological momentary assessment aims to study various phenomena in real-world contexts, minimize recall bias, and maximize ecological validity [12].

This approach has been applied to collect both subjective and objective data from patients with PD in only two preliminary studies, to the best of our knowledge. In the study by van Gilst et al [13], objective metrics of motor performance were collected using an electronic test, while a self-administered pen-and-paper motor diary was used to record subjective symptoms. In a feasibility study by Bot et al [14], only objective metrics of motor features were prospectively collected.

To ensure that patient symptoms are collected prospectively, self-assessed information needs to be recorded at specific time points during the day and integrated with objective assessment in a holistic way. Electronic and information technology devices are thus essential for ecological momentary assessment. However, digital technologies can be impractical for PD subjects, who often have impaired hand dexterity.

Although software apps for smartphones exist for PD [15,16], their use can be challenging for the patients, mainly because of the small screen size on these devices. We are not aware of any tablet-based applications designed for patients with PD that take into account the ergonomic issues specific to this population.

Several systems have been developed to date to collect information on mobility and motor symptoms directly in the home of patients with PD [14-20]. These home-based systems are designed to collect objective [14-20] or subjective [14,19] metrics of motor features. However, to the best of our knowledge, there are currently no available software applications that prospectively record both objective and subjective metrics of motor and nonmotor features of PD.

To meet the need for a tool capable of collecting both objective and subjective metrics of motor and nonmotor symptoms that was handy for PD patients, our group developed an application for tablets called SleepFit. This new application, designed specifically for patients with PD, combines subjective motor scales, a sleepiness scale, and a sleep diary. SleepFit proposes questions and tasks to the patients at specific time points during the day and attributes an exact timestamp indicating when the data are collected. The data collected with SleepFit are automatically stored in a remote server and can be retrieved, integrated with data from other sources, and managed by means of a web application, the SleepFit Researcher Portal.

We aim here to introduce this home-based monitoring system and the improvements that were made in v2.0, the most recent version discussed in the paper. In addition, we outline specific

requirements for software applications targeted to patients with PD. We then share the lessons learned through an iterative development centered on the real needs of patients with PD in real-life conditions. Finally, we evaluate patient compliance and satisfaction with this new tool.

Methods

Participants

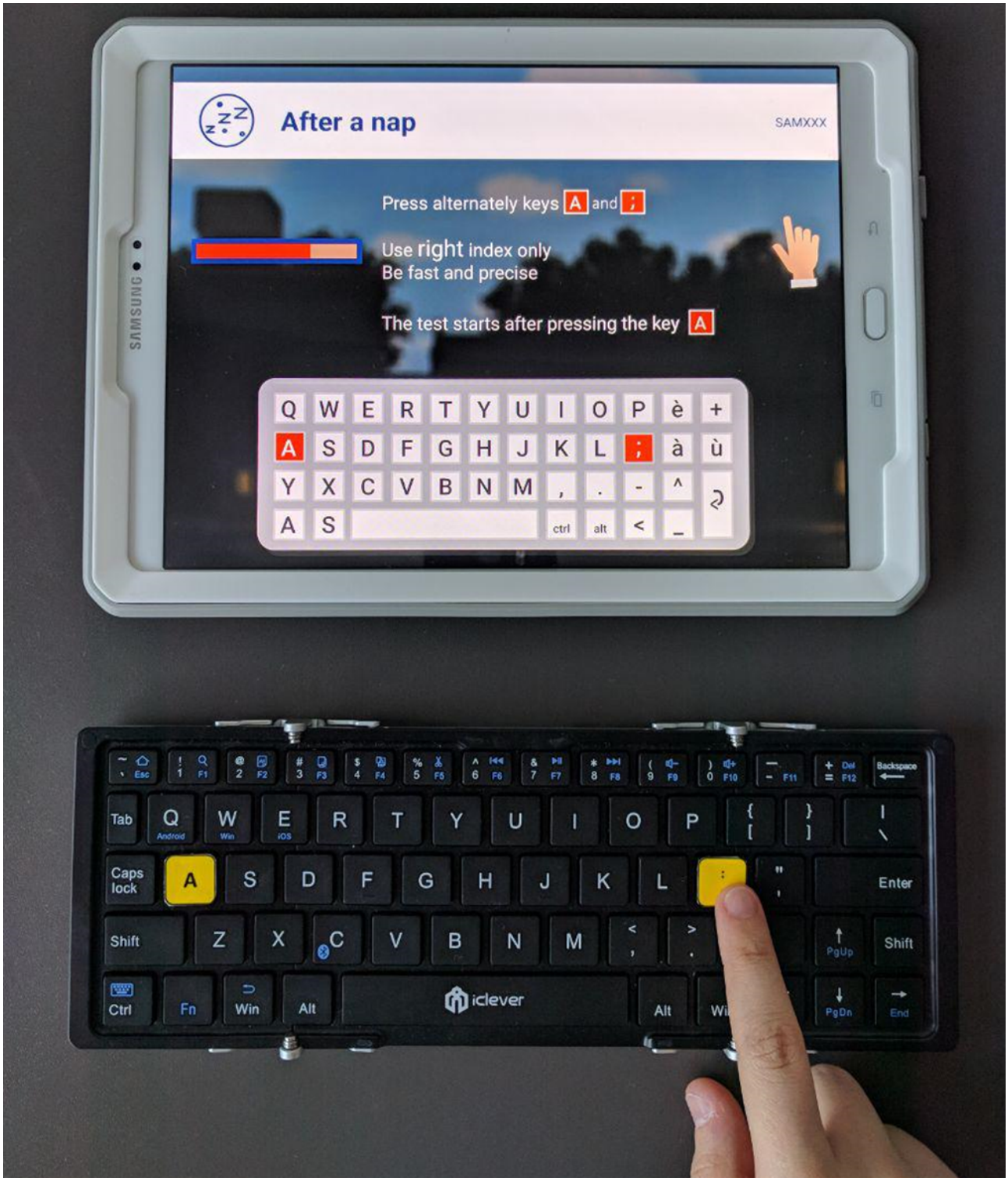
This study was conducted on the first 56 consecutive participants enrolled in the Sleep & Move study between March 2016 and December 2018. Participation in the study was proposed to all consecutive patients meeting the eligibility criteria who were attending the outpatient department of the Movement Disorder Unit of the Neurocenter of Southern Switzerland in Lugano, Switzerland. Other patients volunteered to participate after advertisement of the study in the magazine of the Swiss Parkinson patients' association and in public conferences organized by the same association. Eligibility criteria were mild to moderate idiopathic PD (no atypical parkinsonism) [21] (Hoehn & Yahr stage >1 and ≤3) [22], no cognitive impairment (Mini-Mental State Examination score ≥26/30) [23], no active depression (Beck Depression Inventory score <14/63) [24], no deep brain stimulation.

The initial visit was performed at the hospital by a senior neurologist (PLR) expert in sleep medicine and movement disorders. Each patient underwent a full general clinical and neurological examination including the Movement Disorders Society Unified Parkinson's Disease Rating Scale with the motor part (III) performed during the "on" phase in patients with motor fluctuations. During the initial visit, each patient was given instructions for the use of the SleepFit application. The 14-day home period started at the end of the initial visit for each patient. A follow-up visit was performed at the hospital at day 14 by the same person. The satisfaction questionnaires were collected during this follow-up visit. No economic compensation was provided to the participants. Each participant was offered a reimbursement of the travel expenses for the two visits.

Sleep & Move Study

This usability study was conducted in the framework of the Sleep & Move study [25], which focused on the characterization of sleep benefit in PD. A second objective of Sleep & Move was to test the usability and acceptability of the SleepFit application by patients with PD and add improvements based on patient experience and feedback. Participation in this study was on a voluntary basis. The study was conducted in accordance with the Declaration of Helsinki and written consent was provided by all participants. The study was reviewed and approved by the Ethical Committee of the Canton of Ticino (PB_2016-00056) and registered at ClinicalTrials.gov [NCT02723396] before recruitment began. All participants were asked to use the SleepFit home-based system for 14 consecutive days. The system is composed of an Android tablet on which the SleepFit application is installed and a portable wireless keyboard used for a specific motor task (Figure 1).

Figure 1. The SleepFit home-based system with application installed on a tablet (top) and portable wireless keyboard used to perform motor task (eg, after napping; bottom).



Sessions

Four sessions daily were to be completed by the patients: the "on waking" session was scheduled 30 minutes after waking up in the morning; "after medications", 1 hour after the intake of the first dose of dopaminergic medications in the morning; "afternoon", in the afternoon before taking dopaminergic medication (where applicable); and finally, "evening", just before bedtime.

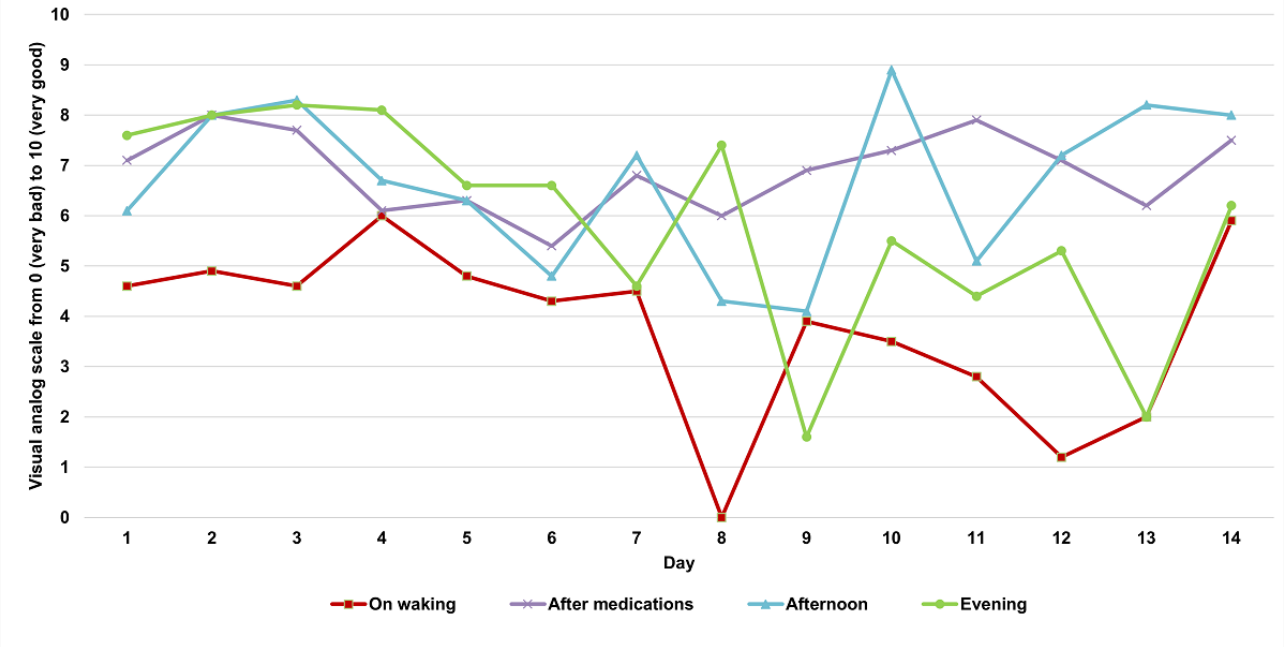
The timing of these 4 daily sessions was chosen considering several disease-related and practical considerations. Four time points a day is the optimum to reliably characterize the daily profile of fluctuation of PD motor or nonmotor symptoms without excessively interfering with the patients' daily routines. Moreover, these 4 moments of the day are the most adapted, from a clinical standpoint, to seize both spontaneous fluctuations of PD and the ones related to dopaminergic medication intake and end-of-dose effects. This format is also in accordance with

routine clinical practice for patients in movement disorder units. An additional session ("nap") could be performed after any nap taken by the patient during the day at his or her request.

Specific Tasks

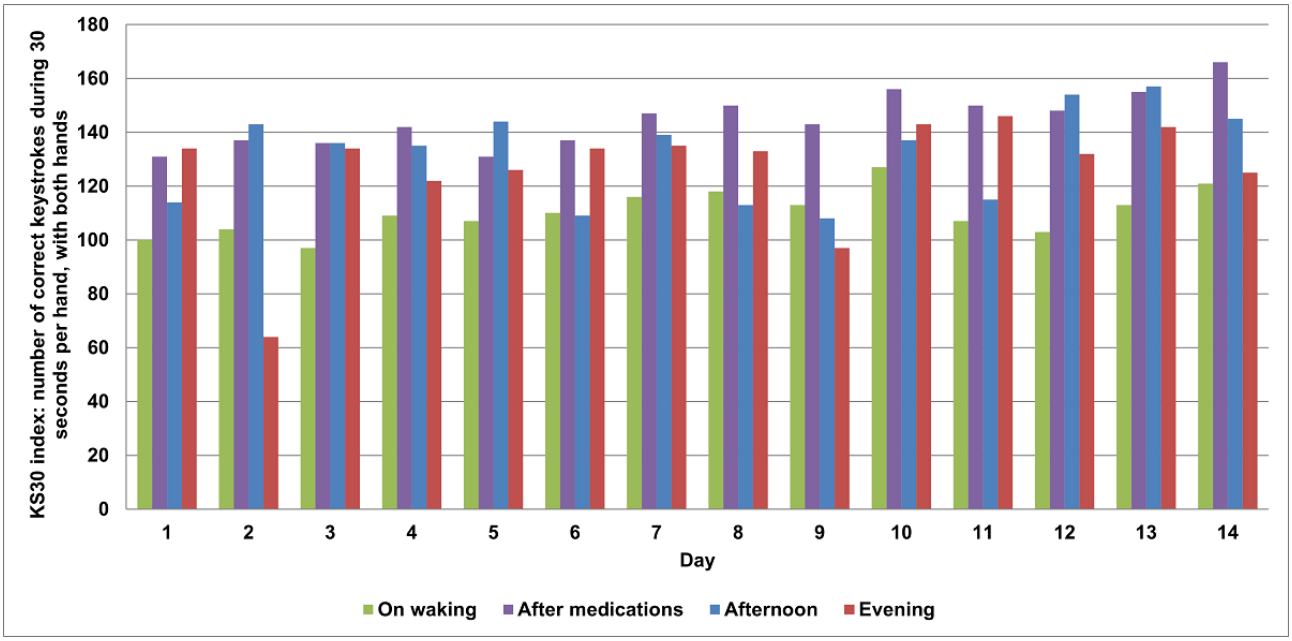
In each session, the patient was asked to perform 2 sets of tasks: subjective scales and a motor test. Subjective scales measure motor status, sleepiness, and emotional state. They include questions 1, 2, 3, and 4 of the Scales for Outcome in Parkinson's Disease Diary Card [26], a visual analog scale (score 0 to 10) [10,27] for perceived overall mobility, tremor, mood, and anxiety level (one scale for each), and the Stanford sleepiness scale [28]. An example of the 14-day patient answer evolution is shown in Figure 2.

Figure 2. Answers to "How would you rate your capability to move right now?" collected on SleepFit for a single patient.



The motor test consists in a digital finger-tapping test that we named the Fit Test, which is based on the previously published Bradykinesia-Akinesia Incoordination test [29,30]. In this test, the subject is asked to strike 2 keys on an external keyboard at 10 cm from each other repeatedly and alternatively for 30 seconds using one hand at a time (Figure 1). We computed the same parameters validated in Noyce et al [29] (see Figure 3 for an example of a parameter computed for a single subject).

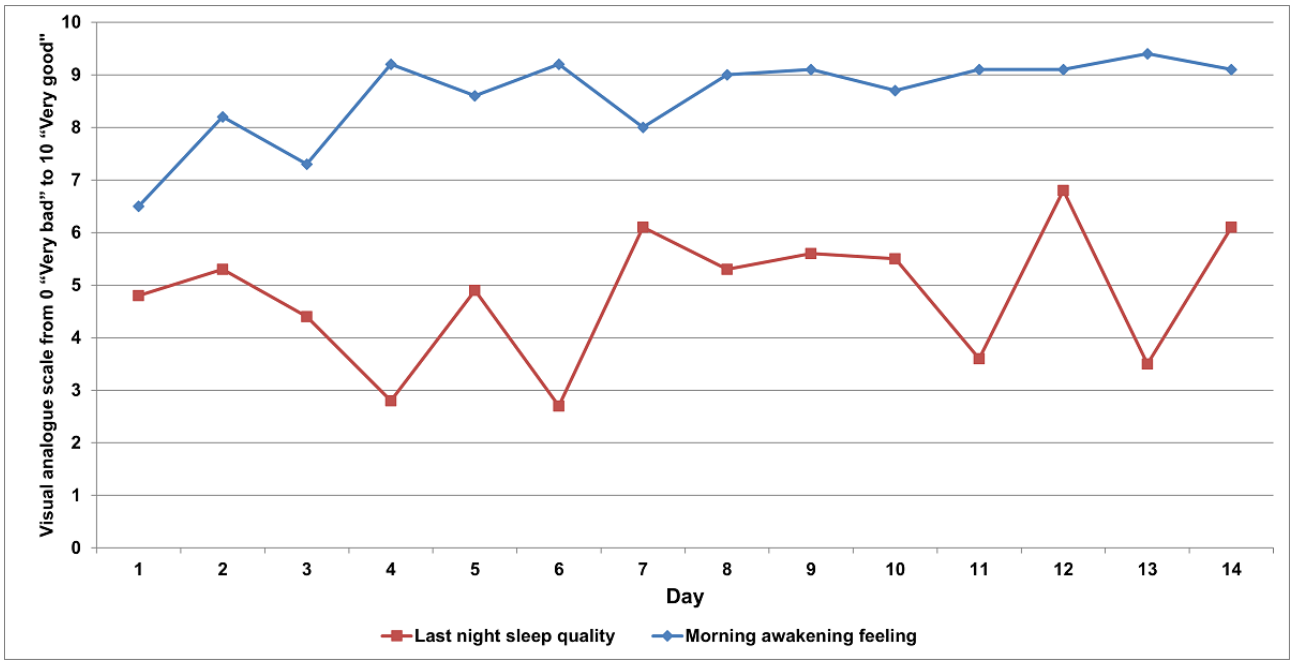
Figure 3. Objective data collected during the execution of a Fit Test by a single patient.



Finally, a 24-question sleep diary was completed once a day during any of the 4 daily sessions. If the patient did not complete the sleep diary during the after medications session, a prompt would appear at each of the following sessions on the same day. The sleep diary includes a characterization of day-to-day sleep habits; use of stimulating or sleep-promoting beverages,

substances, and drugs; and PD-specific or nonspecific nocturnal and diurnal sleep-related symptoms. An example of sleep diary data is shown in Figure 4. Only fully completed tasks (subjective scales, Fit Test, and sleep diary) are recorded and stored by SleepFit.

Figure 4. Answers to "Overall, I slept..." (line with squares) and "This morning, upon awakening, I feel..." (line with diamonds) collected from the sleep diary for a single patient.



SleepFit System

The purpose of SleepFit is twofold: to provide PD patients with a handy home-based system (SleepFit tablet application) suitable for completing the clinical tasks described above (subjective scales, Fit Test, and sleep diary) and provide physicians and researchers with a tool (SleepFit Researcher Portal) to access and consult the data recorded by the application to integrate them with other data sources in the framework of clinical research projects and retrieve them all easily at any time. The modular structure of the SleepFit system enables easy integration of new clinical tasks, functionalities, and external sensors. SleepFit is meant to be shared with other research groups upon mutual agreement.

SleepFit Application and Patient Interface

SleepFit user interface was developed with an iterative and user-centric strategy to take into account the ergonomic issues specific to this population. We opted for a physical keyboard instead of using the in-application touchscreen keyboard since the motor task was validated on a physical keyboard as well. In addition, the following aspects were also considered in favor of the external keyboard: (1) patients included in our study are mainly elderly subjects who may be more familiar with physical keyboards rather than touchscreens; (2) touchscreen keyboards may raise issues related to skin conductivity (eg, sweaty hands) that could prevent correct key touches from being recorded; and (3) patients may experience a better tactile or sensory feedback and more confidence when dealing with physical buttons. Three versions of the application have been released and tested on patients: alpha, v1.0, and v2.0. The relevant differences and enhancements between versions are described in Table 1.

Table 1. Feature improvements between main versions of SleepFit.

Features	Alpha version	Version 1.0	Version 2.0
Graphics	Static background, plain graphics	Dynamic contextual pictograms and wallpapers, session, and task progress bar	Same as v1.0 + font size adapted to the content
Ergonomics	Numerical answers inserted by virtual keyboard	One question per page layout, large-size buttons, predefined numerical answers, +/– buttons for numerical entries	Same as v1.0
Navigation	Session selection relying on the patient	Automatic presentation of sessions, customizable texts and default values, reminders	Same as v1.0 + kiosk mode, multiple language support
Data storage	Remote database (internet connection required)	Store and forward (local + remote database)	Same as v1.0
Remote control	Direct access to remote database	Built-in web application for data querying	Same as v1.0

The alpha version was conceived as a basic software application for tablets bringing together several *ad hoc* tests, scales, and questionnaires employed in clinical routine. Ideas on how to implement the application from the alpha version to v1.0 and subsequently v2.0 arose after direct clinical observation of patients interacting with the application during initial and follow-up visits, comments we received from the participants, and analysis of the more common mistakes they tended to make while using the application interface.

From the prototype (alpha) version to v1.0, the main improvements involved graphics, ergonomics, and navigation. While the layout of alpha version was similar to a paper questionnaire, with plain background, multiple questions on the same page, and standard navigational buttons, graphical and ergonomic improvements in v1.0 included (1) larger, well separated, and accurately positioned buttons to avoid common keypress errors by the participants; (2) ad hoc contextual pictograms and wallpapers; (3) 1 question per page layout; (4) customizable texts, questions, and reminders; and (5) avoidance of open-text fields, which often cause PD patients some difficulty. Regarding navigation, selection of the session to be performed at given times of the day relied on patient choice in the alpha version. From v1.0, a workflow setup guided the patient along the 4 daily sessions by subsequently proposing the correct task to be performed. Additionally, as of v1.0, SleepFit included a log feature, making it possible to record the patient's in-application behavior (clicks on buttons, touches with corresponding x and y screen coordinates, and timestamp of each action taken by the patient). Furthermore, SleepFit v1.0 was equipped with a local database, which ensured that data was saved even in the absence of an internet connection.

Several questions were added to each session of the subjective scales of v1.0 and retained in v2.0 in order to better describe clinically meaningful features that might influence momentary motor performance or its perception such as mood and fatigue, based on clinical observation and patient feedback. This resulted in increased workload for the patients from alpha to v1.0 and v2.0.

Based on patient experience and feedback using SleepFit v1.0, further implementations were included in the September 2018 release of v2.0. In addition to a fine-tuning of ergonomics, the main upgrade to v2.0 was the so-called kiosk-mode feature, which restricts the use of the tablet to the SleepFit application, thus preventing users from accidentally exiting the application and hypothetically interrupting the study. Examples of the SleepFit patient interface are shown in Figure 5.

Figure 5. Screenshots of the SleepFit app: A and B) example of sleep diary questions during “on waking” session; C) execution of subjective scales task during “after medications” session; D) choosing between “afternoon” session and performing tasks before taking medication (“nap” session can be enabled by clicking nap button); E) execution of subjective scales task during “evening session”; and F) bedtime recording during "evening" session (patient clicks “I’m going to bed” just before retiring).

SleepFit Researcher Portal

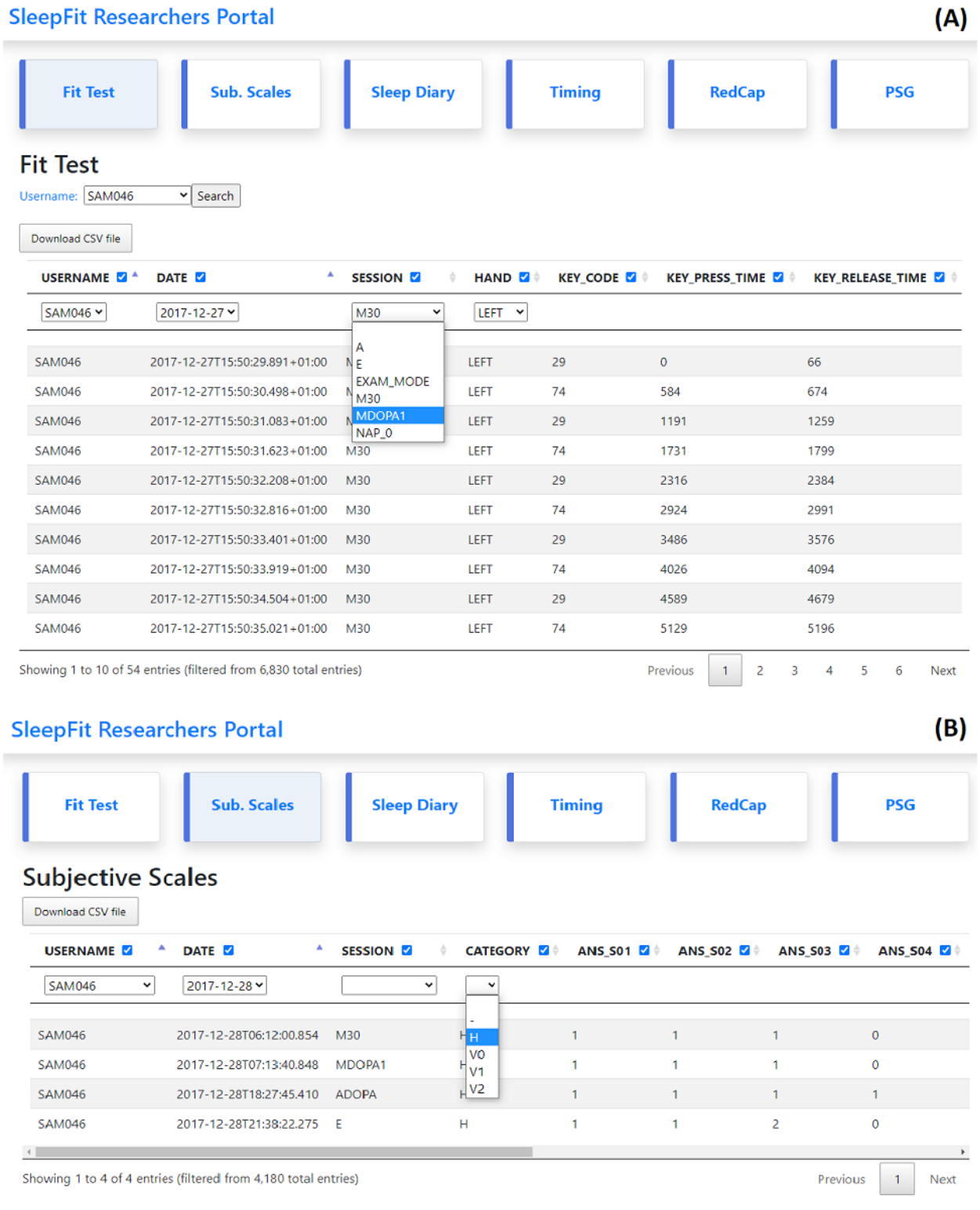
The SleepFit Researcher Portal was conceived and implemented in v1.0 to retrieve and download the data acquired from patients using SleepFit. This portal automatically organizes each patient’s data in sessions and tasks, enabling remote viewing of data from multiple patients. Via the portal, researchers can monitor patient compliance in real time, apply different custom filters (patient ID, date, daily session, etc) to call up specific data, and download data in comma-separated values (CSV) format.

In v2.0, the Researcher portal was further developed as a handy all-in-one tool for data management to be employed in clinical studies using SleepFit. The portal now allows automatic

synchronizing and integrating in the study database of the data collected in the case report form generated by the Research Electronic Data Capture application [31,32], uploading of data in CSV format from other sources or devices, or performing intermediate analyses.

A screenshot of the portal is presented in Figure 6. Each button on the action bar opens a different window to consult, filter, or download data from different sources and of different types: Fit Test, subjective scales, sleep diary, timestamps of bedtime and wake times of the 14 days at home, electronic case report form, and intermediate analyses from video-polysomnographic recordings. The researcher can easily filter the data by selecting the patient ID, date, session, and the hand the test was performed with. The filters patient username, date, session, role, and data categorization (H=data collected at home, V0=data collected at the initial visit; V1=data collected at the first follow-up visit) can be applied, and filtered data can be downloaded in CSV format.

Figure 6. Overview of the SleepFit Researcher Portal for an individual patient: A) Fit Test and B) subjective scales.



Participant Workload

The study was designed in line with the main objective of the Sleep & Move study, which was to systematically characterize potential spontaneous variations in mobility in relation to sleep. We considered that 9 tasks grouped into 4 sessions per day would achieve a good compromise between acceptable patient workload and sufficiently detailed characterization of mobility and sleep symptoms over a 24-hour period. Each patient was therefore asked to complete 9 tasks per day for 14 days: Fit Test (4 times per day), subjective scales (4 times per day), and sleep diary (once per day).

These 9 tasks were split into the 4 daily sessions according to the Sleep & Move study protocol. Patients executed the tasks at specific times during the day. If a patient did not complete a specific session task within a certain maximum time (which varied based on the session itself), the session expired, and that task could not be performed anymore during that day. In the context of the subjective scales and sleep diary, the alpha version proposed 49 and v1.0 and v2.0 proposed 68 questions per day to each patient.

We included in the analyses the tasks performed from the "evening session" of the first day to the "after medications" session of the 14th day. By the end of the home evaluation period, each patient was expected to have completed 124 tasks (55 subjective scales, 55 Fit Test, 14 sleep diaries), answering a total of 680 questions (alpha version) or 941 questions (v1.0 and v2.0). On-demand nap sessions (performed only if the patients had napped and activated the session request themselves) were not included in the analyses as they could not be scheduled a *priori*.

Statistical Analysis

Statistical analyses were performed using the R statistics package (R Foundation for Statistical Computing) [33] and Python programming language (Python Software Foundation) [34]. Patient compliance was calculated on the total of the participants and for each of the 3 versions of SleepFit. Two additional analyses were performed on a subgroup of patients having used v1.0 and v2.0.

Demographic and Clinical Characteristics

To investigate possible differences in the demographic and clinical characteristics of groups of patients having used different versions of SleepFit (between alpha and v1.0 and between v1.0 and v2.0), we assumed samples coming from a normal distribution and used a *t* test for unequal variances for numeric variables and a proportion test for percentages.

Compliance

To assess patient compliance, we computed the ratio of the total number of tasks completed by each patient to the total number of tasks proposed by the application during the home period, as detailed above. The compliance rate is expressed in percentage. The improvement from alpha to v1.0 and v2.0 is evaluated with a classical 1-tailed proportion test.

Familiarity With SleepFit Interface

Patient familiarity with the application was computed by analyzing behavior during application use and keeping track of all user interactions with the touchscreen of the tablet, independently of whether the patient clicks on a zone of the screen associated with a command or not. To do this, we compared the number of hit targets (ie, clicks on the parts of the screen where buttons are located) with the total number of touches on the screen. The total screen touches were tracked thanks to the log feature included as of v1.0 of SleepFit. The target ratio, calculated as the ratio between these 2 values, provides an estimate of how accurate the patients were during application use and how well the user interface has been designed. For average target ratio calculation, we excluded the evening session of the first day and the 2 morning sessions (on waking and after medications) of the 14th day because we included only fully completed days in this analysis. The average target ratio was calculated for each patient and further averaged among all patients in order to provide an overall accuracy value.

Satisfaction

At the end of each subject's participation, a survey regarding patient satisfaction with the home-based study using SleepFit v1.0 and v2.0 was administered (see Multimedia Appendix 1). In this analysis we included 7 questions. Three questions focused on user-friendliness of the application: (1) "Did you encounter difficulties using the SleepFit application?" (2) "Did you encounter difficulties in understanding what to do in the different situations proposed?" (3) "Did you encounter difficulties when inputting the answers to the questions proposed (button selection, timing and quantities, sliders)?" Three other questions assessed the graphical interface of the application: (1) "Were texts clearly legible?" (2) "Were the answer and navigation buttons clear?" (3) "Were wallpapers, pictograms, and colors useful for understanding what to do in each different situation proposed?" One final general question assessed the patient's potential willingness to use the SleepFit application again in the future: "Would you like to use the SleepFit application again in the future if your neurologist proposed it for your clinical follow-up?" All answers were categorical, ranging from 1 (low satisfaction) to 4 (high satisfaction).

Results

Participants

Of the 56 patients included, 51 completed all study procedures; 1 patient dropped out on day 2 because of an inability to use the tablet (alpha version of SleepFit) and 4 patients prematurely terminated their participation because they perceived an excessive burden due to the study protocol (n=3) or because of physical ailments (n=1). Analyses were conducted on 52 patients (14 females), corresponding to 92.9% of the initial population. The first 19 participants tested the alpha version of SleepFit, 7 used v1.0, and 26 used v2.0. Figure 7 depicts the participant flow. Demographic and clinical characteristics of the study population are shown in Table 2.

Figure 7. Participant flow.

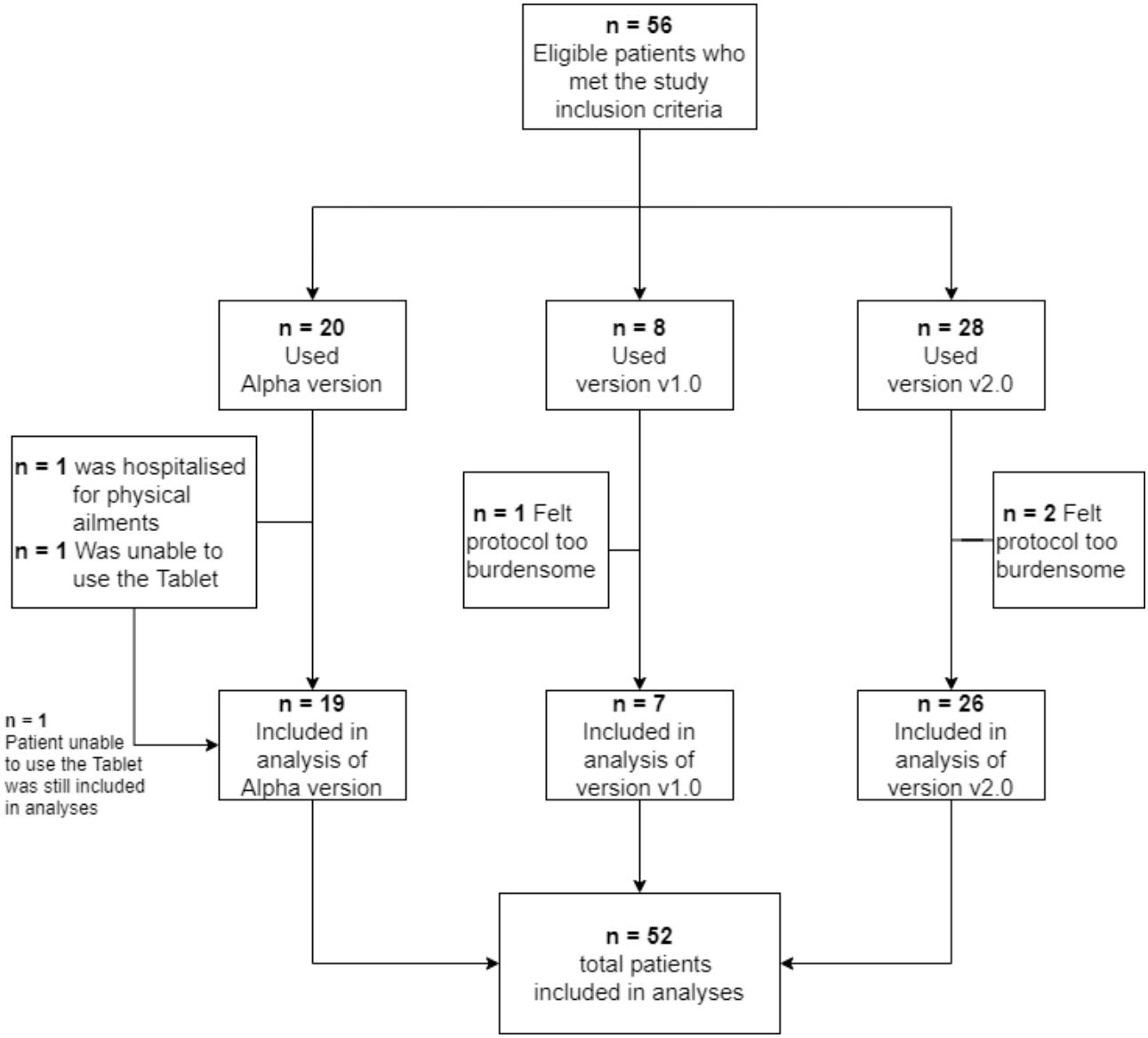


Table 2. Patient demographic and clinical characteristics for each version of SleepFit and all versions together.

Demographics	Alpha version (n=19)	Version 1.0 (n=7)	Version 2.0 (n=26)	All versions (n=52)
Age (years), mean (SD)	61.9 (9.5)	68.9 (11.5)	66.7 (9.8)	67.8 (9.8)
Sex (female), n (%)	3 (16)	1 (14)	10 (39)	14(27)
Active worker, n (%)	3 (16)	2 (29)	8 (31)	13 (25)
Smartphone user, n (%)	12 (63)	6 (86)	22 (85)	40 (77)
Attempts to independently use SleepFit, mean (SD)	2.0 (1.3)	1.4 (0.5)	1.5 (0.5)	1.7 (0.9)
Parkinson disease duration (years)	5.7 (3.1)	7.8 (7.8)	7.9 (6.5)	7.1 (5.7)
Hoehn & Yahr stage	2.1 (0.2)	1.8 (0.4)	1.9 (0.5)	2.0 (0.4)
MDS-UPDRS[a] I	8.2 (3.6)	10.6 (5.0)	8.9 (4.8)	8.9 (4.4)
MDS-UPDRS II and IV	10.2 (5.7)	13.7 (8.2)	12.3 (7.7)	11.7 (7.1)
MDS-UPDRS III	37.3 (12.7)	26.1 (9.9)	27.3 (12.4)	30.8 (13.0)
Levodopa daily equivalent dose (mg)	507.7 (271.0)	561.5 (476.9)	666.2 (447.4)	594.2 (395.4)
CIRS-G[b] total score	8.7 (4.2)	5.4 (4.2)	6.5 (4.2)	7.0 (4.3)
Pittsburgh Sleep Quality Index	6.1 (3.1)	6.4 (3.2)	6.8 (2.8)	6.5 (2.9)
Fatigue Severity Scale	3.9 (1.4)	5.0 (1.7)	4.3 (1.5)	4.3 (1.5)

[a]MDS-UPDRS: Movement Disorder Society–Sponsored Revision of the Unified Parkinson's Disease Rating Scale.

[b]CIRS-G: Cumulative Illness Rating Scale–Geriatric.

General Design and Development Principles

The implementation of SleepFit brought us to some general considerations which might be helpful for the conception and development of software applications dedicated to patients with PD:

- Tablet format seems to be particularly suitable for software applications to be used by patients with PD
- When different sessions are required to be performed during the day, the software should guide patients through them automatically to avoid mistakes
- Presenting 1 question per page helps patients focusing their attention on the question being asked
- Ergonomics should be considered carefully: large-size buttons, predefined answers, avoiding free-texts answers, simplified navigation to improve usability
- When researchers or clinicians directly provide hardware support, kiosk-mode results are particularly advantageous for those patients not familiar with information technology
- Store-and-forward technology to a cloud-based database helps minimize data loss
- A web platform for data querying allows the investigators/clinicians to provide the patients with remote assistance during the ongoing data collection
- A tool to integrate, synchronize, and retrieve all study data, such as the SleepFit Researcher Portal, might be very useful for data management
- In order to meet the needs of physicians, the structure of the system should be modular and customizable, allowing easy integration of new clinical tests, functionalities, or external sensors

Demographic and Clinical Characteristics

Demographic and clinical characteristics of each group of patients using different versions of SleepFit are summarized in Table 2. No significant difference among groups was found except for PD duration, which showed a difference between v1.0 and v2.0 (P=.03); we interpret this difference to be due to the small size of the sample.

Compliance

The total expected number of tasks of all patients taken together was 6420. Our analysis revealed that 88.89% (5707/6420) were effectively completed. Detailed estimates of compliance for each version of SleepFit are reported in Table 3.

Considering all 9 daily tasks together, the average daily workload for each patient was 8 minutes and 41 seconds for the alpha version of SleepFit, and 11 minutes and 15 seconds for v1.0 and v2.0. The last 2 versions of the application required more time because the number of questions on the subjective scale task increased from 5 (alpha) to 11 (v1.0 and v2.0) to assess additional motor and sleep-related symptoms. The average total time spent by each patient performing study procedures over the duration of the study was 2 hours and 35 minutes.

Table 3. Patient compliance for each version of SleepFit and all versions together. Compliance is reported as the percentage of the tasks completed by the patients over the total of proposed tasks.

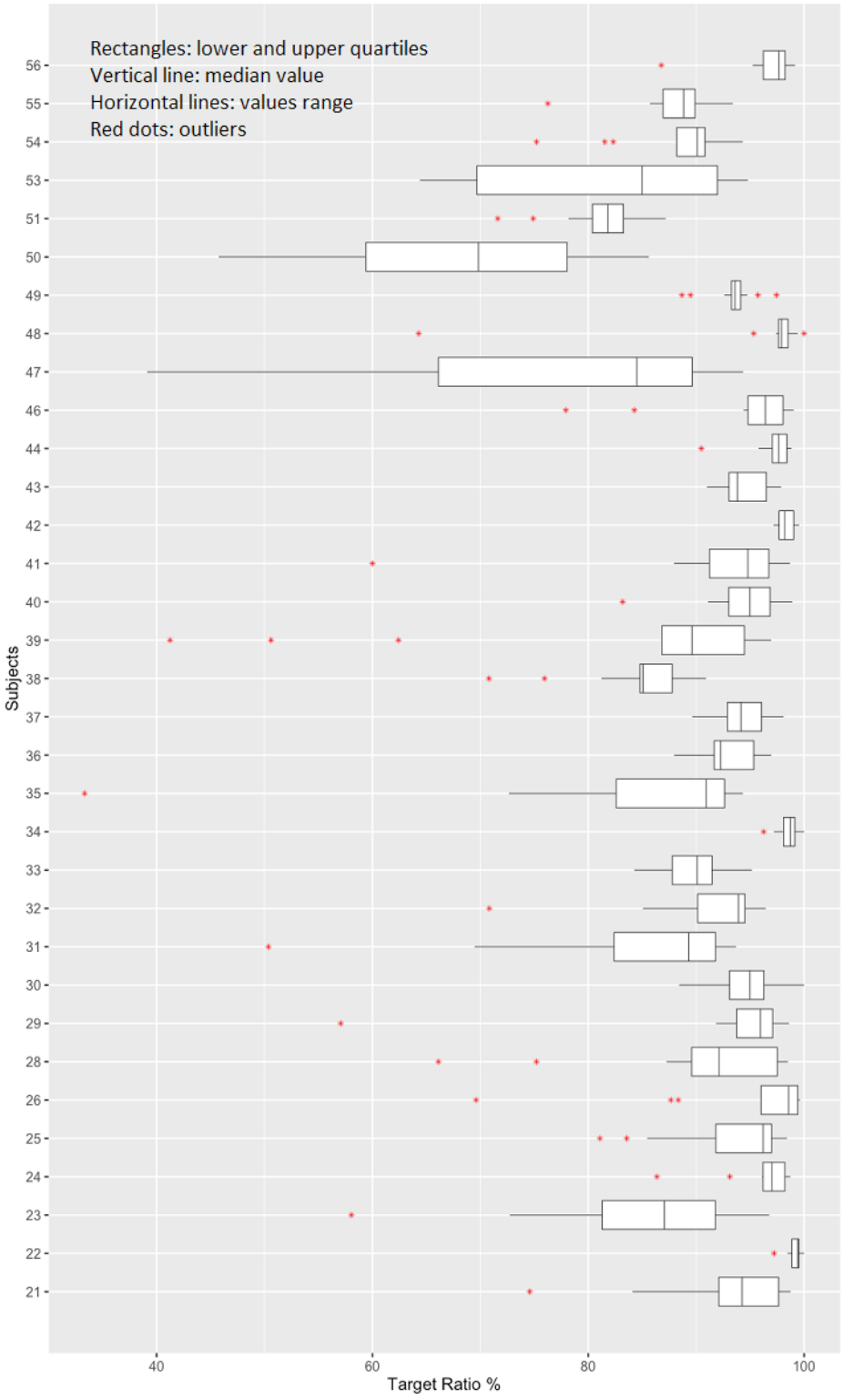
Characteristic	Alpha version (n=19) 680 questions in 14 days: 8 mins per day			Version 1.0 (n=7) 941 questions in 14 days: 11 mins per day			Version 2.0 (n=26) 941 questions in 14 days: 11 mins per day			All versions (n=52)		
	Subscale (%)	Fit test (%)	Sleep diary (%)	Subscale (%)	Fit test (%)	Sleep diary (%)	Subscale (%)	Fit test (%)	Sleep diary (%)	Subscale (%)	Fit test (%)	Sleep diary (%)
On waking	88.7	90.6	—[a]	88.8	93.9	—	88.2	89.6	—	88.5	90.5	—
After medications	87.2	86.1	—	88.1	89.3	—	86	86.8	—	86.7	86.8	—
Afternoon	87.5	86.2	—	73.6	82.4	—	92.9	92.9	—	88.3	89.1	—
Evening	86.8	—	—	87.8	88.8	—	86.3	87.4	—	86.7	86.3	—
Total per test	87.6	86.7	93.6	84.6	88.7	96.7	88.3	89.1	99.7	87.5	88.2	97.1

[a]Not applicable.

Familiarity With SleepFit Interface

The first subscale analysis concerning patient behavior during application use was performed on the 33 patients who used v1.0 and v2.0 of SleepFit (ie, the versions equipped with the log feature). This analysis revealed that, on average, the target ratio across all patients was 90.1 (SD 10.4; range 33.3-100). The average target ratio during the home evaluation period for each subject is shown in Figure 8.

Figure 8. Average target ratio over the home period for the 33 subjects who used SleepFit v1.0 and v2.0.



Satisfaction

According to the satisfaction questionnaires (completed by the 26 patients who used v2.0 of SleepFit only), 96% (25/26) would use SleepFit again for clinical purposes for the same period (17/26, 65%) or even for longer (8/26, 31%). Only one patient declared that he would not use it again. For what concerns the usability aspect, 71% (55/78) declared no difficulty, 24% (19/78) slight difficulty, 4% (3/78) moderate difficulty, and 1% (1/78) high difficulty. Finally, the patients rated the graphical interface as follows: 92% (72/78) optimal, 7% (5/78) good, and 1% (1/78) could be better; no one rated the interface very poor.

Discussion

Principal Findings

To the best of our knowledge, SleepFit is the first tablet application and remote monitoring system specially designed

to be employed in real-life settings for patients with PD that collects subjective and objective clinical data on motor and nonmotor symptoms and subjective sleep information. SleepFit went through several phases of development before reaching v2.0. This innovative, iterative, and user-centric approach targeted to patients with PD is the result of a very close and fruitful collaboration among clinicians, patients, information technologists, and biomedical and telecommunication engineers.

A home-based system could be perceived as intrusive by patients with PD in their daily routines and, thus, may not always be well accepted. Analyzing patient compliance provides a measure of how well these systems are accepted and integrated into the everyday life of patients. We assume this provides an estimate of expected compliance in clinical practice and research too. Compared with the data presented in the literature [35], SleepFit represents a valid alternative to pen-and-paper questionnaires. In fact, it shows higher overall compliance. This result is all the more noteworthy considering that, compared with the study by Stone et al [35], our population includes patients with PD and older age individuals who may experience greater difficulties in dealing with technological devices compared with younger, healthy people. Moreover, the workload required for patients in our study is somewhat greater than in previous studies to date [14-16,20], of which, however, only one evaluated compliance and satisfaction with the system when used on a strictly scheduled protocol [14]. Indeed, our patients answered 68 questions and performed 4 Fit Test tasks every day, for a total of 941 questions and 55 Fit Test tasks by the end of the 14th day. Patients were also asked to respect the time constraints, which could have made it harder for them to adhere to the study since they performed the 4 daily sessions at specific times of the day and had a maximum time allowed for the execution of each task. These aspects may have negatively influenced patient compliance. In spite of this, we observed higher compliance on similar workloads and time constraints compared with studies that used similar technological tools [16].

According to our analysis, there was a significant (P=.04) increase of about 2% in compliance between patients who used v1.0 and v2.0 of SleepFit. The sleep diary was completed with high compliance, at 99.7%, in SleepFit v2.0. This value may appear extremely high at first glance but is actually plausible and accurate. The reason is that the sleep diary questionnaire was proposed to the patient at each session throughout the day until it was completed. This means that there were no time constraints on completing the sleep diary, and therefore, better compliance is expected. Conversely, the decrease in compliance during the day from the first morning session to the evening session was not negligible. This may suggest that incremental fatigue accumulating during the daytime may have negatively influenced patient motivation to perform the tasks.

The data collected from the satisfaction questionnaires also provided useful information about patient acceptance of the SleepFit application. Patient willingness to use SleepFit again seems to indicate that this tool might be suitable for clinical follow-up of patients and also for research studies requiring longer participation periods. In fact, almost 31% of the patients who completed the survey reported that they would be willing to use SleepFit again, even for a longer period. The target ratio analysis performed on 33 subjects provided a more objective measure of the usability of SleepFit. We found that the patients were satisfactorily accurate when using the application, with more than 90% of screen touches done on target locations of the screen. This suggests that the design of the graphical interface of SleepFit is suitable for patients with impaired movement capabilities and tremor issues.

Limitations

We did not distinguish in our analysis if noncompleted tasks were due to technical bugs or a true lack of compliance by the patient. Therefore, our analysis may have underestimated actual compliance, since all missing data were treated as not provided by the patients. Finally, a further limitation of this study is that the conditions in which the tests were performed at home were not verified. However, this limitation is intrinsic in all home-based assessment methods and could only be overcome by means of an increased patient burden (eg, if a researcher came to the patients' homes during each test session), which would seriously hamper the utility of a home-based approach.

Future Implementations

Future developments of SleepFit will include data quality checks based on the use of synchronized accelerometric data from wearable sensors. We also foresee the possibility of integrating other tasks and metrics such as electronic tests to assess cognition or other subjective scales exploring motor or nonmotor symptoms of PD. Due to the modular structure of SleepFit, adding new tests or scales is straightforward. SleepFit can thus be easily adapted to different research protocols and for clinical use. The results of our study provide valuable information on the possibility of using SleepFit in other contexts. For instance, a 2-week study paradigm including tests and assessments performed 4 times a day appears to be well accepted by PD patients. SleepFit might also be suitable for clinical follow-up of patients living in remote areas, for chronic patients during the confinement imposed by health authorities in response to the COVID-19 pandemic, or for people with mild cognitive dysfunction, thanks to its user-friendliness.

Conclusion

SleepFit is an easy-to-use tool that can accurately collect subjective and objective data from patients with PD. It was developed and improved with an iterative user-centric approach. From the lessons learned in this process, essential suggestions emerged for future software application development tailored to PD patients. Although the use of SleepFit should be further tested in larger populations, both for clinical follow-up and in other home-based research studies, this application proved to be a very promising tool to increase patient compliance and assist researchers in surveying patients during data collection and data management.

Acknowledgments

The development of SleepFit v1.0 received the contribution of the DOS Group (Mr Raffaele Calabrese and Mr Piero Bozzolo). We would like to thank all the patients who took part in this study. Many thanks to Dr Stephany Fulda for her support in the first conception of the Sleep, Awake & Move research project and to Miss Simona Bonoli and Miss Lucia Guglielmetti for their help in patient recruitment. A special thanks to Mr Gregor Ruggel for the German translation of the SleepFit texts and Miss Aurélie Davin for linguistic revision of the French version. We also acknowledge the invaluable contribution of Prof Moustapha Dramé for his advice on the methodological revision of the final manuscript, and we really appreciate the linguistical revision made by Mrs Fiona Ecarnot. We also thank all the colleagues and coworkers who supported this project and shared our enthusiasm: Dr Ninfa Amato, Dr Salvatore Galati, Dr Claudio Staedler, Dr Filippo Scacchi, Mr Marco Bosetti, Mr Luigi Fiorillo, Mr Pierluigi Lurà, Mrs Nicole Vago-Caputo, Mr Matteo Pereno, Mrs Simona Perrotta, Mr Giuliano Filippini, and Mrs Yasmin Belloni.

Authors' Contributions

EKC and YL conceived and developed the alpha version of SleepFit; v1.0.0 and following were conceived by AP, MM, PLR, AM with the support of AKL and CF and developed by AM.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Satisfaction questionnaire for the SleepFit app.
[DOCX File , 41 KB - mhealth_v9i6e16304_app1.docx]

References

1. de Rijk M, Launer L, Berger K, Breteler M, Dartigues J, Baldereschi M, et al. Prevalence of Parkinson's disease in Europe: a collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology 2000;54(11 Suppl 5):S21-S23. [Medline: 10854357]
2. Zhang Z, Román G. Worldwide occurrence of Parkinson's disease: an updated review. Neuroepidemiology 1993;12(4):195-208. [doi: 10.1159/000110318] [Medline: 8272179]
3. Karlsen KH, Larsen JP, Tandberg E, Maland J. Quality of life measurements in patients with Parkinson's disease: a community-based study. Eur J Neurol 1998 Sep;5(5):443-450. [doi: 10.1046/j.1468-1331.1998.550443.x] [Medline: 10210872]
4. Oerlemans WG, de Weerd AW. The prevalence of sleep disorders in patients with Parkinson's disease. A self-reported, community-based survey. Sleep Med 2002 Mar;3(2):147-149. [doi: 10.1016/s1389-9457(01)00127-7] [Medline: 14592234]
5. Giannoccaro MP, Antelmi E, Plazzi G. Sleep and movement disorders. Curr Opin Neurol 2013;26(4):428-434. [doi: 10.1097/wco.0b013e3283632cef]
6. Barone P, Antonini A, Colosimo C, Marconi R, Morgante L, Avarello TP, PRIAMO study group. The PRIAMO study: a multicenter assessment of nonmotor symptoms and their impact on quality of life in Parkinson's disease. Mov Disord 2009 Aug 15;24(11):1641-1649. [doi: 10.1002/mds.22643] [Medline: 19514014]
7. Avidan A, Hays R, Diaz N, Bordelon Y, Thompson A, Vassar S, et al. Associations of sleep disturbance symptoms with health-related quality of life in Parkinson's disease. J Neuropsychiatry Clin Neurosci 2013;25(4):319-326 [FREE Full text] [doi: 10.1176/appi.neuropsych.12070175] [Medline: 24247858]
8. van Gilst MM, Bloem BR, Overeem S. "Sleep benefit" in Parkinson's disease: a systematic review. Parkinsonism Relat Disord 2013 Jul;19(7):654-659. [doi: 10.1016/j.parkreldis.2013.03.014] [Medline: 23615667]
9. Aarsland D, Brønnick K, Larsen JP, Tysnes OB, Alves G, Norwegian ParkWest Study Group. Cognitive impairment in incident, untreated Parkinson disease: the Norwegian ParkWest study. Neurology 2009 Mar 31;72(13):1121-1126. [doi: 10.1212/01.wnl.0000338632.00552.cb] [Medline: 19020293]
10. Ratti P, Faraci F, Hackethal S, Mascheroni A, Ferlito C, Caverzasio S, et al. A new prospective, home-based monitoring of motor symptoms in Parkinson's disease. J Parkinsons Dis 2019;9(4):803-809 [FREE Full text] [doi: 10.3233/JPD-191662] [Medline: 31476169]
11. Olanow CW, Stern MB, Sethi K. The scientific and clinical basis for the treatment of Parkinson disease (2009). Neurology 2009 May 26;72(21 Suppl 4):S1-S136. [doi: 10.1212/WNL.0b013e3181a1d44c] [Medline: 19470958]
12. Moskowitz D, Young S. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20 [FREE Full text] [Medline: 16496031]
13. van Gilst M, van Mierlo P, Bloem B, Overeem S. Quantitative motor performance and sleep benefit in parkinson disease. Sleep 2015 Oct 01;38(10):1567-1573 [FREE Full text] [doi: 10.5665/sleep.5048] [Medline: 25902811]
14. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]

15. Arora S, Venkataraman V, Zhan A, Donohue S, Biglan K, Dorsey E, et al. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: a pilot study. Parkinsonism Relat Disord 2015 Jun;21(6):650-653. [doi: 10.1016/j.parkreldis.2015.02.026] [Medline: 25819808]
16. Arora S, Baig F, Lo C, Barber TR, Lawton MA, Zhan A, et al. Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. Neurology 2018 Oct 16;91(16):e1528-e1538 [FREE Full text] [doi: 10.1212/WNL.0000000000006366] [Medline: 30232246]
17. Heldman D, Urrea-Mendoza E, Lovera L, Schmerler D, Garcia X, Mohammad M, et al. App-based bradykinesia tasks for clinic and home assessment in parkinson's disease: reliability and responsiveness. J Parkinsons Dis 2017;7(4):741-747. [doi: 10.3233/JPD-171159] [Medline: 28922169]
18. Pan D, Dhall R, Lieberman A, Petitti DB. A mobile cloud-based Parkinson's disease assessment system for home-based monitoring. JMIR Mhealth Uhealth 2015 Mar 26;3(1):e29 [FREE Full text] [doi: 10.2196/mhealth.3956] [Medline: 25830687]
19. Mera TO, Heldman DA, Espay AJ, Payne M, Giuffrida JP. Feasibility of home-based automated Parkinson's disease motor assessment. J Neurosci Methods 2012 Jan 15;203(1):152-156 [FREE Full text] [doi: 10.1016/j.jneumeth.2011.09.019] [Medline: 21978487]
20. Lee CY, Kang SJ, Hong S, Ma H, Lee U, Kim YJ. A validation study of a smartphone-based finger tapping application for quantitative assessment of bradykinesia in parkinson's disease. PLoS One 2016;11(7):e0158852 [FREE Full text] [doi: 10.1371/journal.pone.0158852] [Medline: 27467066]
21. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992 Mar;55(3):181-184 [FREE Full text] [doi: 10.1136/jnnp.55.3.181] [Medline: 1564476]
22. Hoehn M, Yahr M. Parkinsonism: onset, progression and mortality. Neurology 1967 May;17(5):427-442. [doi: 10.1212/wnl.17.5.427] [Medline: 6067254]
23. Folstein MF, Robins L, Helzer J. The Mini-Mental State Examination. Arch Gen Psychiatry 1983 Jul;40(7):812. [doi: 10.1001/archpsyc.1983.01790060110016] [Medline: 6860082]
24. Beck AT. An inventory for measuring depression. Arch Gen Psychiatry 1961 Jun;4:561-571. [doi: 10.1001/archpsyc.1961.01710120031004] [Medline: 13688369]
25. ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT02723396 [accessed 2021-05-13]
26. Marinus J, Visser M, Stiggelbout AM, Rabey JM, Bonuccelli U, Kraus PH, et al. Activity-based diary for Parkinson's disease. Clin Neuropharmacol 2002;25(1):43-50. [doi: 10.1097/00002826-200201000-00008] [Medline: 11852296]
27. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990 Aug;13(4):227-236. [doi: 10.1002/nur.4770130405] [Medline: 2197679]
28. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology 1973 Jul;10(4):431-436. [doi: 10.1111/j.1469-8986.1973.tb00801.x] [Medline: 4719486]
29. Noyce A, Nagy A, Acharya S, Hadavi S, Bestwick J, Fearnley J, et al. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. PLoS One 2014;9(4):e96260 [FREE Full text] [doi: 10.1371/journal.pone.0096260] [Medline: 24781810]
30. Giovannoni G, van Schalkwyk J, Fritz VU, Lees AJ. Bradykinesia akinesia inco-ordination test (BRAIN TEST): an objective computerised assessment of upper limb motor function. J Neurol Neurosurg Psychiatry 1999 Nov;67(5):624-629 [FREE Full text] [doi: 10.1136/jnnp.67.5.624] [Medline: 10519869]
31. 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] [Medline: 18929686]
32. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
33. R: a language and environment for statistical computing. 2013. URL: http://www.R-project.org [accessed 2021-05-13]
34. Python Foundation. URL: https://docs.python.org/3.6/reference/, [accessed 2021-05-13]
35. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials 2003 Apr;24(2):182-199. [doi: 10.1016/s0197-2456(02)00320-3] [Medline: 12689739]

Abbreviations

CSV: comma-separated values
PD: Parkinson disease

Edited by L Buis; submitted 17.09.19; peer-reviewed by P Markopoulos, D Pan, D O'Dell, B Cole; comments to author 10.08.20; revised version received 31.10.20; accepted 26.03.21; published 08.06.21.

Please cite as:
Mascheroni A, Choe EK, Luo Y, Marazza M, Ferlito C, Caverzasio S, Mezzanotte F, Kaelin-Lang A, Faraci F, Puiatti A, Ratti PL
The SleepFit Tablet Application for Home-Based Clinical Data Collection in Parkinson Disease: User-Centric Development and Usability Study
JMIR Mhealth Uhealth 2021;9(6):e16304
URL: https://mhealth.jmir.org/2021/6/e16304
doi:10.2196/16304
PMID:34100767

Review

Challenges With Developing Secure Mobile Health Applications: Systematic Review

Bakheet Aljedaani[1,2], MSc; M Ali Babar[1,3], PhD

[1]Centre for Research on Engineering Software Technologies, School of Computer Science, The University of Adelaide, Adelaide, Australia

[2]Computer Science Department, Aljumum University College, Umm Alqura University, Makkah, Saudi Arabia

[3]Cyber Security Cooperative Research Centre, Adelaide, Australia

Corresponding Author:
Bakheet Aljedaani, MSc
Centre for Research on Engineering Software Technologies
School of Computer Science
The University of Adelaide
North Terrace
Adelaide,
Australia
Phone: 61 883135208
Email: bakheet.aljedaani@adelaide.edu.au

Abstract

Background: Mobile health (mHealth) apps have gained significant popularity over the last few years due to their tremendous benefits, such as lowering health care costs and increasing patient awareness. However, the sensitivity of health care data makes the security of mHealth apps a serious concern. Poor security practices and lack of security knowledge on the developers' side can cause several vulnerabilities in mHealth apps.

Objective: In this review paper, we aimed to identify and analyze the reported challenges concerning security that developers of mHealth apps face. Additionally, our study aimed to develop a conceptual framework with the challenges for developing secure apps faced by mHealth app development organizations. The knowledge of such challenges can help to reduce the risk of developing insecure mHealth apps.

Methods: We followed the systematic literature review method for this review. We selected studies that were published between January 2008 and October 2020 since the major app stores launched in 2008. We selected 32 primary studies using predefined criteria and used a thematic analysis method for analyzing the extracted data.

Results: Of the 1867 articles obtained, 32 were included in this review based on the predefined criteria. We identified 9 challenges that can affect the development of secure mHealth apps. These challenges include lack of security guidelines and regulations for developing secure mHealth apps (20/32, 63%), developers' lack of knowledge and expertise for secure mHealth app development (18/32, 56%), lack of stakeholders' involvement during mHealth app development (6/32, 19%), no/little developer attention towards the security of mHealth apps (5/32, 16%), lack of resources for developing a secure mHealth app (4/32, 13%), project constraints during the mHealth app development process (4/32, 13%), lack of security testing during mHealth app development (4/32, 13%), developers' lack of motivation and ethical considerations (3/32, 9%), and lack of security experts' engagement during mHealth app development (2/32, 6%). Based on our analysis, we have presented a conceptual framework that highlights the correlation between the identified challenges.

Conclusions: While mHealth app development organizations might overlook security, we conclude that our findings can help them to identify the weaknesses and improve their security practices. Similarly, mHealth app developers can identify the challenges they face to develop mHealth apps that do not pose security risks for users. Our review is a step towards providing insights into the development of secure mHealth apps. Our proposed conceptual framework can act as a practice guideline for practitioners to enhance secure mHealth app development.

(JMIR Mhealth Uhealth 2021;9(6):e15654) doi:10.2196/15654

KEYWORDS

systematic literature review; mHealth apps; secure apps; developers; security knowledge

Introduction

Background

The use of mobile apps in health care has gained widespread adoption [1,2]. Lack of health professionals, especially in rural areas, is an excellent motivator for mobile health (mHealth) app adoption [3]. mHealth apps rely on the portability and context-sensitivity of mobile computing to improve access to health care services that are cost-effective, scalable, and pervasive [4]. Leveraging mHealth apps would improve access to health care services, lower the cost, and increase patients' health awareness [5]. According to the World Health Organization, mHealth is defined 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" [6]. There are several types of mHealth apps developed for health purposes ranging from general health apps such as decision, support, vitals, and reproductive health apps through fitness apps providing an activity tracker, nutrition tracker, and mindfulness [5]. The number of mHealth apps has grown massively following the launch of centralized mobile app repositories (ie, Google Play and Apple Store) in 2008. It has become easier for mobile developers to distribute their apps to a wide range of users [7]. Research2guidance, an organization for providing research and consultancy for digital health, reports that 78,000 new mHealth apps were added to apps stores in 2017. The report also showed that mHealth app downloads reached 3.7 billion, and the market revenue for digital health reached US $5.4 billion in 2017 [8].

The security of mobile apps in general and mHealth apps in particular has become one of the primary concerns since mobile apps are more vulnerable to attacks [9]. Most mobile apps collect, process, store, and transmit user and device data in and out of a device over various networks [5]. Compromising the confidentiality, integrity, and availability of such data would lead to severe consequences, including but not limited to compromised device data and leading to financial loss [10]. In mHealth apps, security becomes a significant concern due to health-critical data privacy and integrity [5,11]. An attack to falsify clinical measurements can lead to unnecessary care for patients as they think they are sicker than they actually are and can cause medical, legal, and social concerns [12].

Health professionals are increasingly relying upon health data collected via mHealth apps to make their decisions, such as dermatologic care [13], chronic illnesses management [14,15], and clinical practice [16]. Data manipulation can significantly impact treatment, causing serious results (eg, worsened morbidity or death) [17,18]. While health regulations and laws (ie, The US Health Insurance Portability and Accountability Act [HIPAA], European General Data Protection Regulation [GDPR]) strive to protect medical integrity and patients' privacy by focusing on hospitals, doctors, and insurance firms, little attention has been paid to support mHealth app developers by providing them with suitable guidelines for developing secure apps [5,19].

A large part of mHealth app security relies on developers' experience with designing and developing secure apps. We use the term developer in our research to refer to professionals who are engaged in engineering and development of mHealth apps. According to previous studies [1,12,15,20,21], most mHealth apps have not fully implemented mechanisms to protect health data. Studies have also claimed that mHealth developers may fail to appropriately implement basic security solutions such as authentication, encryption for data at rest, and encryption for data in transit. It is being recognized that it is critically important to thoroughly train mHealth app developers in implementing suitable security mechanisms to protect patients' data from being stolen or compromised [12,20]. Hence, it is crucial to identify and synthesize the reported challenges of developing secure mHealth apps as a body of knowledge for research and practice. We have reviewed the relevant literature to determine the security challenges by focusing on developers rather than the solutions. Our research question for this literature review is: What are the challenges that developers of mHealth apps face with respect to implementing security?

This review's primary contribution is identifying the challenges that hinder the development of secure mHealth apps, such as the lack of security guidelines and regulations for developing secure mHealth apps and developers' lack of knowledge of and expertise with secure mHealth app development. This review's results can be beneficial to researchers and practitioners (eg, mHealth app developers, managers, research engineers) for supporting research and development of emerging and next-generation, secure mHealth apps.

Previous Work

The challenges for developing secure software have been receiving increasing attention in recent years. A review by Kanniah and Mahrin [22], which included 44 studies, identified the factors that influence secure software development practices. The study found that security skills, expertise, tools, and development time are among the factors that impact secure software development. The identified factors were classified into institutional context, people and action, project content, and software development process factors. Thomas et al [23] addressed the issues that security auditors face during application review for security bugs. The study recommend further support for the development process by providing security-related tools and effective communication tools for developer interactions. Further support for software developers has also been recommended by providing motivation (eg, reward or recognition) and providing solutions for technical challenges such as using third-party library issues. The authors recommended recruiting security experts within teams and make them available for answering questions. Raghavan et al [24] presented a model for achieving security during the software development lifecycle (SDLC). Their model suggests the following factors: security policy, management support, security-related training for developers, and development process control. Weir et al [25] studied the positive factors that enhance the development of secure software. The work identified the interventions that lead to achieving security by performing a threat model, organizing motivational workshops to engage team members, and continuous reminders for developers. The study also highlighted other interventions that need to be considered, such as component choice for security

tools, performing static analysis, developer training, and performing penetration testing and code review.

Some studies also aimed to help mobile app developers develop secure apps by providing guidelines for the development process [6,26,27]. Given the increasing realization of the need to provide developers of mHealth apps with appropriate knowledge, training, and support for developing secure apps, there is a critical need to identify and analyze the challenges that prevent them from developing secure apps. Our findings would contribute to a body of knowledge about the challenges that mHealth app developers face with respect to security.

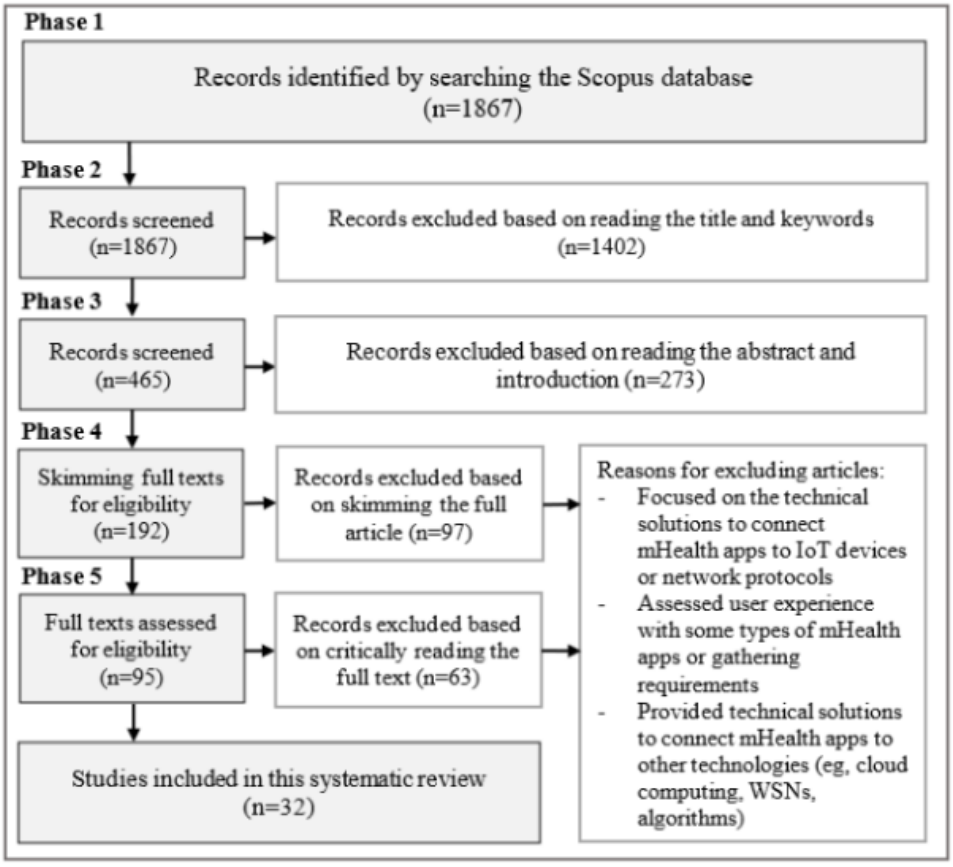
Comparison With Prior Studies

Prior reviews [28,29] have focused more on investigating the security measures and technical solutions employed by developers. However, a few challenges were raised in [28,29]. Katusiime and Pinkwart [28] systematically reviewed privacy and usability issues and solutions in mHealth systems. The study considered developers' lack of security knowledge and lack of a security framework as external factors that need to be considered. Another review by Marquez et al [29] was more on the security issues of telehealth systems. The study focused on classifying security (ie, attacks, vulnerabilities, weaknesses, and threats) and presenting security strategies (ie, detect attacks, stop or mitigate attacks, and react to attacks) of telehealth systems. Also, the study reported some security practices that need to be ensured, such as having a discussion about architectural styles (eg, security patterns) and engaging stakeholders during the development of an app. To the best of our knowledge, there is no systematic literature review (SLR) that explicitly investigates the challenges faced by mHealth app developers when implementing security for mHealth apps. Thus, we aimed to fill the gap and provide insights into the development of secure mHealth apps.

Methods

This research has been undertaken as an SLR. It is one of the most widely used research methods of evidence-based software engineering. An SLR provides a well-defined process for identifying, evaluating, and interpreting all available evidence relevant to particular research. We followed the guidelines of Kitchenham et al [30] to perform an SLR that involves 3 main phases: defining a review protocol, conducting the review, and reporting the review. In this section, we briefly describe the main components of the review protocol and its implementation. Our review protocol has 6 components, including research question, search strategy, data source, study selection process, inclusion and exclusion criteria, and data extraction and data synthesis. Figure 1 presents a flow diagram of the literature search and article selection results.

Figure 1. Flow diagram for the selection of articles. IoT: Internet of Things; mHealth: mobile health; WSN: wireless sensor network.



Research Question

Our review's objective was to identify and codify the challenges that hinder mHealth app developers from developing secure apps. This review's findings would enable us to identify the potential gaps that need to be further investigated based on the developers' perspectives.

Search Strategy

We used the following strategies to form our search string: (1) identifying the major terms based on the study focus and the research question, (2) identifying all the possible keywords and related synonyms based on our experience and previous work, (3) using the Boolean "AND" to join major terms and the Boolean "OR" to join alternative terms and synonyms. Hence, our search string for this review was as follows: ("security" OR "insecure" OR "secure") AND ("mobile health" OR "mobile healthcare" OR "mobile health-care" OR "mobile health care" OR "telehealth" OR "mhealth").

Data Source

We used the Scopus digital library as our primary search library as there are many successful examples of other researchers (eg, [30]) limiting their search to Scopus. The Scopus indexing system has the advantages of facilitating the formulated complex search string, being frequently updated, and keeping track of a large number of journals and conferences in software engineering studies. Furthermore, Scopus is an indexing database that provides name, keywords, and abstract for all published articles. Any pointed articles can be further searched and downloaded to review the whole article regardless of which database in which it actually exists.

Study Selection Process

As illustrated in Figure 1, we followed several criteria to exclude studies in our SLR as detailed in the following sections.

Phase 1: Automatic Search

We ran our search string in the Scopus digital library. Thus, we retrieved a total of 1867 potential articles.

Phase 2: Title and Keyword-Based Selection

We carefully reviewed the title and keywords to decide whether each of the retrieved articles was relevant to our SLR. We retained the papers for the next inspection when we could not decide by reading the titles and keywords. Thus, we excluded 1402 articles and included 465 articles for the next phase.

Phase 3: Abstract and Introduction-Based Selection

We read the abstract and introduction for each article. This phase enabled us to include 192 articles and discard 273 articles.

Phase 4: Full Paper Scanning–Based Selection

We scanned the entire article to ensure that it was relevant to our SLR objective. Thus, we included 95 articles and excluded 97 articles.

Phase 5: Critical Review–Based Selection

We critically reviewed the included papers and excluded duplicates (eg, extended versions of the studies were included, and shorter versions were excluded). Thus, we excluded 63 articles and included 32 studies, referred to as S1 to S32. A list of the included papers is presented in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

For the purpose of this review, we applied predefined inclusion and exclusion criteria for paper selection. We included primary studies that focused on the development process of secure mHealth apps, studies written in English published from January 2008 to October 2020 since major app stores (Google Play and Apple Store) were launched in 2008, and peer-reviewed publications (ie, journals, conferences, workshops, and book chapters).

Besides excluding non-peer-reviewed studies (ie, lecture notes, summaries, panels, and posters) and studies that were not written in English, we excluded studies that contained irrelevant content for our review such as studies that focused on investigating technical solutions (eg, encryption methods, authentication mechanisms, access control) for mHealth apps; studies providing technical solutions to connect mHealth apps to Internet of Things (IoT) devices or cloud computing technology; studies that focused on sensor layers (eg, wireless sensor networks), developing algorithms, or network protocols for mHealth apps; studies that focused on mHealth app quality or gathering functional requirements; and studies that examined user experiences with some mHealth apps (eg, patient management apps).

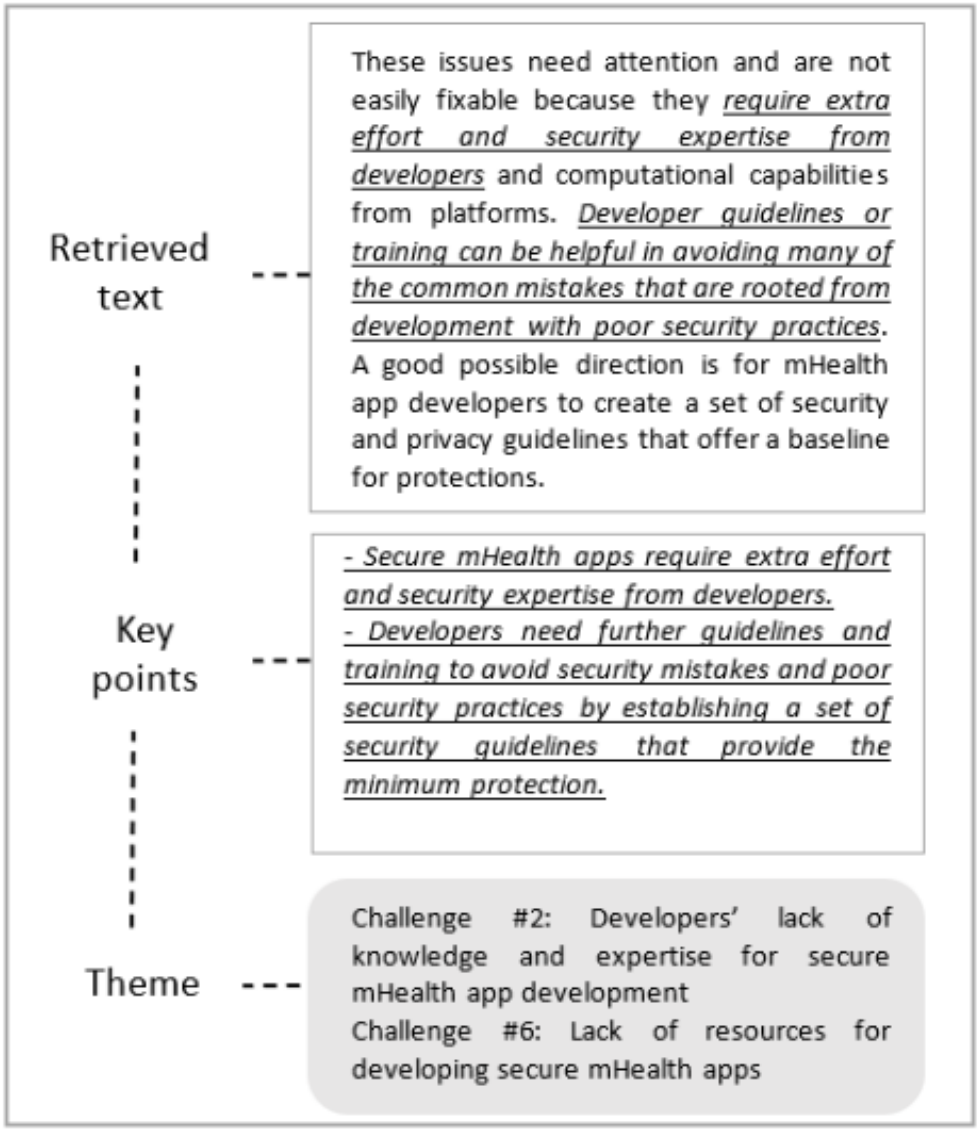
Data Extraction and Synthesis

We divided the extracted data into 2 categories: study characteristics and the challenges for developing secure mHealth apps. Our data extraction form is shown in Multimedia Appendix 2. We performed descriptive statistics to analyze the demographic data. To answer our research question, we used the Endnote tool to manage the bibliography and utilized Excel spreadsheets to extract and synthesize the data. We used thematic analysis, a qualitative analysis technique, to analyze and synthesize the extracted data to derive the results for this review [31]. We mainly followed the thematic analysis method's 5 steps: (1) familiarizing oneself with the data, which involved trying to read and examine the extracted data items; (2) generating initial codes, which involved extracting the initial lists of challenges; (3) searching for themes, which involved trying to combine different initial codes generated from the second step into potential themes; (4) reviewing and refining themes, which involved checking the identified challenges from step 3 against each other to understand what themes had to be merged with others or dropped; and (5) defining and naming themes, which involved defining a name for each challenge. Figure 2 demonstrates an example that was taken from S4 [32] of how our final list of challenges was identified.

To further enhance our analysis, we developed a conceptual framework to present the correlation among the identified challenges. We followed the steps of Regoniel [33] to develop a conceptual framework that involves 4 steps: choose the topic, do a literature review, isolate the important variables, and generate the conceptual framework. It should be noted that the initial coding was done by the first author and was reviewed and revised (followed by a discussion wherever required) with

2 independent researchers, Dr Leonardo Iwaya and Dr Faheem Ullah, who are experts in the field of mHealth apps and doing SLR studies to avoid potential bias.

Figure 2. Example of the steps of applying the thematic analysis to the qualitative data. mHealth: mobile health.



Results

We now present the findings of our SLR. We classified the findings into demographic information, challenges for developing secure mHealth apps, and the conceptual framework for the identified challenges.

Study Characteristics

In this subsection, we present the study characteristics based on the outlet (ie, journal, conference, or workshop) of the selected papers, as shown in Table 1.

Providing such information would be helpful for new researchers interested in conducting research in this particular area. We selected 32 primary studies for this review. The complete list of the reviewed articles is available in Multimedia Appendix 1. All selected studies mainly discussed the security aspects of mHealth apps. Table 1 shows the distribution, year of publication, and different outlets. It should be noted that our reviewed studies were published from 2012 to 2020. Of 32 studies, we noticed that 23 studies (72%) were published as journal papers; 7 studies (22%) were published in conferences, while 2 studies (6%) were published as workshop papers. Furthermore, we noticed that 11 studies (34%) were published in JMIR and JMIR mHealth and uHealth, and 2 studies were published at International Conference on Future Internet of Things and Cloud Workshops (2017, 2019).

Table 1. The number of selected studies published per year and their distribution by outlet.

Year	Journals, n	Conferences, n	Workshops, n
2012	1	1	0
2013	0	0	0
2014	5	1	0
2015	4	2	0
2016	2	0	0
2017	3	0	1
2018	4	0	0
2019	4	1	1
2020	0	2	0

Challenges With Developing Secure mHealth Apps

This subsection reports the results based on our analysis to answer the study research question: "What are the challenges that developers of mHealth apps face with respect to implementing security?" Our analysis identified 9 challenges (referred to as C1 to C9) that hinder app developers from developing secure mHealth apps. The identified challenges were ordered based on their frequency within the reviewed studies. Table 2 illustrates the identified challenges, the key points that led us to consider them from the reviewed studies, and the frequency of each challenge.

Table 2. Challenges with developing secure mobile health (mHealth) apps (identified from 32 studies).

Challenge number and description	Key points from reviewed studies	Frequency, n (%)
C1. Lack of security guidelines and regulations for developing secure mHealth apps	Lack of security guidelines, regulations, direct laws about the security requirements, secure designing, security testing, security features that need to be employed in mHealth apps (S4 [32], S5 [17], S6 [34], S7 [35], S10 [36], S12 [9], S13 [37], S15 [38], S16 [19], S20 [39], S22 [20], S23 [1], S26 [40], S29 [41], S31 [42]); lack of framework or standards (eg, standardized policies and methodologies to ensure the security standards are met) for developing secure mHealth apps (S2 [43], S3 [12], S29 [41], S31 [42]); lack of compliance with the available guidance and/or standard (S25 [44], S29 [41]); challenges for the developers to deal with legal obligations, policies, and procedures (S32 [4])	20 (63)
C2. Developers' lack of knowledge of and expertise with secure mHealth app development	Insufficient knowledge of software developers about the security risks of mHealth apps (S12 [9], S17 [45], S18 [46], S27 [47]); lack of developers' security awareness (eg, towards the potential threats of mHealth apps; S3 [12], S9 [11], S14 [18], S21 [15], S28 [48], S32 [4]); developers' lack of knowledge towards secure coding practices, using secure APIs[a], and utilizing up-to-date libraries (S18 [46]) or secure third-party services by mHealth app developers that could misuse users' health data (S1 [21], S11 [5], S19 [2], S24 [49]); developers' lack of knowledge towards utilizing security measures (eg, TLS[b] security for servers, proper protection for user passwords) of mobile devices (S3 [12], S8 [50], S22 [20], S25 [44]); lack of experience in secure software development for developers (S4 [32]); lack of auditing security knowledge and review what knowledge they have (S25 [44])	18 (56)
C3. Lack of stakeholders' involvement during mHealth app development	Lack of stakeholders' participation during the development lifecycle of mHealth apps (S5 [17], S10 [36], S20 [39], S29 [41], S30 [51]); lack of security understanding by health professionals when they engage in the development process causing poor elicitation of security requirements (S5 [17])	6 (19)
C4. No or little attention by developers towards the security of mHealth apps	Developer' assumption that users are not concerned about security (S32 [4]); security is not developers' concern (S11 [5], S21 [15]); security issues should be resolved by the testers (S32 [4]); developers with no security focus skip all security measures (S18 [46]); developers are not considering secure design principles and privacy guidelines (S31 [42])	5 (16)
C5. Lack of financial resources for developing secure mHealth apps	No/low budget assigned for employing security measures (S32 [4]); unavailability of security tools (S32 [4]); developers lack training about developing secure mHealth apps (S4 [32], S5 [17]); lack of research and development efforts to facilitate developing secure mHealth apps (S14 [18])	4 (13)
C6. Time constraints during mHealth app development process	Rushing to market, which leaves vulnerabilities in mHealth apps (S18 [46], S26 [40], S32 [4]); the long process of gaining consent or approving the development choices of the developers (S7 [35])	4 (13%)
C7. Lack of security testing during mHealth app development	Lack of security testing (S32 [4]); lack of proper security testing (eg, vulnerability scan) for mHealth apps (S6 [34], S18 [46], S23 [1])	4 (13)
C8. Developers lack motivation and ethical considerations	Lack of motivations for developers during the development process of mHealth apps (S27 [47]); developers lack ethics during the development process of mHealth apps (S10 [36], S30 [51])	3 (9)
C9. Lack of security experts' engagement during mHealth app development	Lack of collaboration and discussion with security experts from the beginning of the development lifecycle of mHealth apps (S18 [46], S32 [4])	2 (6)

[a]APIs: application programming interfaces.

[b]TLS: transport security layer.

C1: Lack of Security Guidelines and Regulations for Developing Secure mHealth Apps

Security guidelines refer to a set of suggested actions or recommendations for things to do or avoid during software development [52]. Security guidelines help app developers, mostly inexperienced, adopt effective security practices and write secure codes. They contain accessible information, properly layered and searchable, with good coverage of all security aspects (eg, cryptography, handling user input and privileges [26]). It would be ideal to clarify that there are numerous security guidelines for ensuring mobile app security (eg, Open Web Application Security Project [OWASP]). According to Nurgalieva et al [53], the available security guidance for developing secure mHealth apps can be categorized into (1) guidelines, recommendations, or principles; (2) app development practices (ie, applied security mechanisms) to ensure mHealth security; and (3) models of user behavior and preferences related to security or privacy. Such guidelines (eg, GDPR) have had the effect of raising awareness and establishing a minimal set of expectations. However, they do not address the issue of the development of systems that meet privacy and security requirements [53]. Additionally, Assal and Chiasson [54] indicated that security guidelines do not exist or are not mandated by the companies or that developers might lack the ability or proper expertise to identify vulnerabilities despite having general security knowledge. Our reviewed studies, including S3 [12], S4 [32], S12 [9], and S20 [39], have pointed

out a general lack of security guidelines for developing secure mHealth apps. Zubaydi et al [9] called for effective guidelines that can help developers build secure mHealth apps (S12 [9]). Even though there are guidelines to protect health data (ie, HIPAA), they do not provide specific instructions for developing secure mHealth apps. Furthermore, it has also been claimed that there is a lack of security frameworks, standards, compliance checklists, and regulations (S22 [20], S18 [46], S13 [37], S20 [39], S2 [43], S9 [11]). Legal restrictions (ie, obtaining security certification) ensure that mHealth app development organizations are not developing vulnerable mHealth apps (S11 [5], S12 [9]).

C2: Developers' Lack of Knowledge of and Expertise With Secure mHealth App Development

The security knowledge of mobile app developers plays a significant part in developing secure mHealth apps. Lack of security knowledge would result in creating an insecure app that may leak health-critical data to attackers. The reviewed studies indicated that mHealth app developers do not have enough security education covering important security aspects. Consequently, developers follow insecure programming practices (eg, employing improper security solutions; S22 [20], S19 [2], S25 [44]) or improper handling for mHealth app permissions (S23 [1]). Furthermore, developers' lack of security knowledge may lead to incorrect security choices when attaching a particular device with mHealth apps (eg, tracking device that helps monitor user behavior; S11 [5], S12 [9], S18 [46]) or integrating an app with other systems (S13 [37]). Making an incorrect security decision may allow health apps to share health-critical data with other mobile apps, untrusted apps, or external hosts (S12 [9]). mHealth app developers may make security-centric decisions based on their best assumption or strategies (S24 [49]). Thamilarasu and Lakin (S18 [46]) conducted a vulnerability scan that revealed 248 vulnerabilities in the top 15 Android-based mHealth apps. The study revealed that the 3 most common vulnerabilities were not errors in systems, but instead, errors in developers' choices (ie, selecting a suitable cipher, choice of permissions to request on a mobile device). The study concluded that most vulnerabilities could have been prevented through proper coding and secure engineering practices.

Keeping in mind that the threat landscape is changing rapidly, dealing with the volatile environment requires developers to keep their security knowledge sharp. Even security experts need to update their knowledge [55]. Despite the fact that mHealth app vulnerabilities are frequently announced in security-relevant knowledge banks (eg, National Vulnerabilities Database, data breach reports) to advise developers, for some reason (ie, difficult to use), these security alerts are not followed or ignored. As a result, unfixed bugs might allow attackers to perform malicious activities (eg, illegally access health-critical data by exploiting sensor permissions, enabling them to extract data or transfer malware to an app [56]). Announcements of identified security bugs are one way of encouraging mHealth developers to keep up to date with the threat landscape. Müthing et al [2] and Dehling et al [18] indicated that mHealth app developers use out-of-date security measures (S14 [18], S19 [2], S23 [1]). As a result, some mHealth apps even have previously exposed security errors (S23 [1]). Despite the realization of the importance of keeping mHealth developers aware of the latest security issues, there is a little evidence that developers get regular formal security training to maintain their security knowledge (S24 [49]). Lack of auditing among developers to maintain and review their security knowledge can create a knowledge gap and lead to out-of-date security knowledge (S25 [44]).

C3: Lack of Stakeholders' Involvement During mHealth App Development

Involving stakeholders in security requirement engineering is being recognized as key to software success and getting effective and impactful outcomes [22]. Indeed, stakeholders' involvement contributes to the elicitation and specification of security requirements of the developed software. Yet, it is difficult, as developers would first exert significant effort to understand the complexity of a problem domain [57]. Also, more time and resources would be required. For mHealth apps, developers should refer to stakeholders (eg, medics, patients) throughout the development process to ensure that the technology meets their needs (S10 [36], S30 [51]). Further, stakeholders need to be involved earlier in the development process of mHealth apps. However, development practices often include clinicians and experts but more rarely involve the target audience until evaluation (S29 [41]). At the same time, it would be challenging for some stakeholders to have an understanding of security due to their capabilities when engaging them in the development process. As a result, this causes poor elicitation of security requirements (S5 [17], S20 [39]).

C4: Lack of Financial Resources for Developing Secure mHealth Apps

The development process of mHealth apps can be supported by using security resources to enhance secure mHealth app development. Lack of necessary resources, such as technology, is a challenge that can directly impact developing secure mHealth apps. For example, security tools (eg, Zed Attack Proxy, Android Debug Bridge, Codified Security, White Hat Security, and Quick Android Review Kit) are resources to facilitate writing secure code and testing apps during the development process. They help developers catch errors that they might be unaware of and adjust their code accordingly before releasing an app. Wurster and van Oorschot [58] argued that not all software developers are security experts, and there is a need to use suitable security tools during a development project. Security tools for mobile apps have received a lot of attention from researchers. A security tool called FixDroid [59] can show warning messages with recommendations to fix errors during the coding phase. It has proven to be effective in improving the security of the written code, is limited to Android app developers, and is not widely known.

Similarly, software libraries can be used as supportive resources to facilitate the software development process. Such libraries help developers reuse specific code for certain goals and support access to hardware and software that might be needed. Yet, it can be challenging for developers to know which library to trust while developing mHealth apps. There can be a risk of data leakage by using untrusted libraries (S16 [19], S13 [37]). Some

libraries, especially the open-source libraries, may collect data about users without developers being aware of it, leading to data privacy breaches [60]. Furthermore, using untrusted third-party libraries to integrate mHealth apps with electronic health records can result in attackers gaining unauthorized access to patients' data (S13 [37]).

Older versions of security resources (ie, tools and libraries) also contain known vulnerabilities (S18 [46]). Most of the security resources are often updated to address security-related issues and introduce new functions; hence, it is important to be aware of and use the latest security tools and libraries. Therefore, developers' security knowledge of the adopted security resources can significantly impact the developed app's security. Besides being aware of the relevant security resources, it can be difficult for developers to learn to use them within the time and resources available for a project (S25 [44], S17 [45]).

C5: No or Little Attention by Developers Towards the Security of mHealth Apps

Incorporating security should ideally be considered throughout SDLC from requirement analysis to the deployment phase [61]. In fact, addressing security at later stages of app development or after app release in the form of security patches can be a costly exercise and can introduce new vulnerabilities [62]. Studies, such as S11 [5], S21 [15], and S31 [42], found that mHealth app developers pay little or no attention to the security of mHealth apps. This issue can be seen for a few reasons, including (1) developer' assumption that users are not concerned about security, (2) developers' assumption that security should be handled by app testers, and (3) developers with no security focus would even skip all security measures to resolve other quality attributes including usability and performance (S18 [46], S32 [4]). Therefore, it is important to come up with effective mechanisms for overcoming developers' lack of attention towards security.

C6: Time Constraints During mHealth App Development Process

Due to business pressures (eg, rushing to market), delivering an app on time tends to be the main aim mHealth apps developers try to satisfy for customers and avoid extra costs. High workload and tight timeframes require mHealth app developers to put more effort in meeting functional requirements as a primary task (S18 [46], S26 [40], S32 [4]). It also affects their attitude and behavior towards addressing security (eg, underestimating risks, assuming attackers will not realize the weaknesses) and dealing with security after releasing an app [61]. This approach leads to insecure mHealth apps, increases the cost, and introduces new vulnerabilities after fixing the existing vulnerabilities [63]. It is estimated that the cost can be 30 to 100 times more expensive to retrofit security compared with incorporating security from the beginning [64]. Besides, the speed of delivering apps will not allow team members to share and convey security knowledge among mHealth app developers [65]. Furthermore, the long process of gaining consent or manager approval of the developers' choices can be an issue (S7 [35]). As a result, this lengthens the process of getting their opinion on a specific task. Hence, this leads to skipping security issues that need to be fixed.

C7: Lack of Security Testing During mHealth App Development

Security testing is one of the essential phases of the mHealth app development lifecycle. Security testing helps determine the quality of apps by ensuring all the security requirements are met. Security testing for mHealth apps, in particular, will help figure out how an app will react against different attacks (eg, unauthorized access to health data, tampering with health data, or reporting invalid health data to health professionals; S11 [5]). Security testing of mHealth apps can be overlooked since it can be a challenging task for developers. Several factors can affect performing security testing, including the absence of security testing tools, lack of effective and well-known testing guidelines, cost of performing app testing by a third-party organization, or lack of a security expert within a software development organization (S23 [1], S18 [46], S6 [34]). Consequently, this would release mHealth apps without conducting security testing, leaving an app at high risk [66]. Wurster and van Oorschot [58] indicated that security testing is not a first-choice task for developers, and their main job is completing the required features.

C8: Lack of Security Experts' Engagement During mHealth App Development

A security expert, security leader, or security champion within an organization plays a vital role during the mHealth app development process (S7 [35]). Besides the development activities, they direct mHealth app developers on secure development practices and perform a security review to ensure their code does not have security defects. A security expert can encourage developers to achieve security goals and educate other developers about the potential threats and solutions (S14 [18]). Lack of security experts within a software development team can lead to failures in applying proper security controls by mHealth app developers. Besides, the lack of availability of security experts would be a challenge for developers (S7 [35]). As a result, there is a lack of constructive feedback that prevents developers from (1) acquiring security knowledge, (2) gaining hands-on experience, and (3) developing apps that are secure by design.

C9: Developers' Lack of Motivation and Ethical Considerations

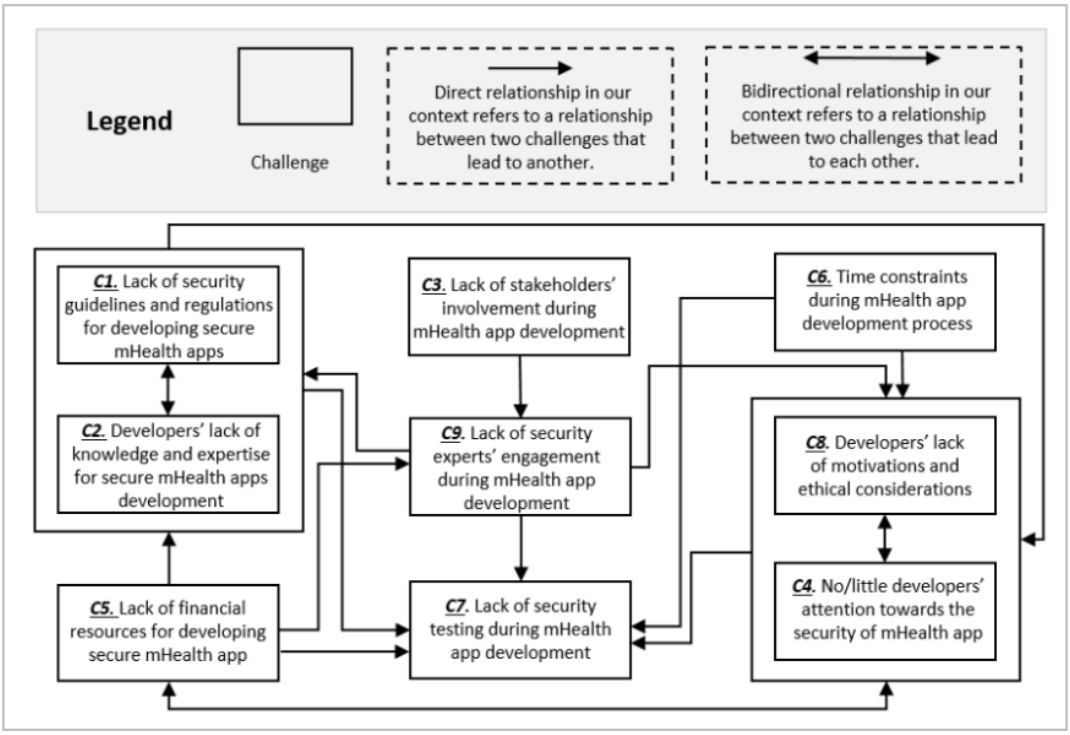
Motivation refers to the driving force behind all the actions of developers during development. It has been recognized as a critical success factor for software projects. Motivation can be seen differently based on developers and an organization's size [67]. The research on security practices indicates that many security incidents are mainly caused by human rather than technical failure [68]. Developers with low motivation were found to be one of the most frequently cited causes of software development project failures [69]. Xie et al [70] presented the reasons that make software developers make security errors. The study concluded that most software developers have a "not my problem" attitude, which indicates that software developers are the source of security errors due to their attitudes and behaviors. In particular, in mHealth app development, studies such as S10 [36], S27 [47], and S30 [51] reported that

developers' lack of motivation and ethical considerations is a challenge that hinders developing secure mHealth apps.

Conceptual Framework

Based on our analysis of the extracted data, we propose a conceptual framework, as in Figure 3, that represents the challenges for developing secure mHealth apps. Jabareen [71] defined a conceptual framework as "a network, or a plane of interlinked concepts that together provide a comprehensive understanding of a phenomenon or phenomena." Figure 3 presents a conceptual framework for correlating the identified challenges.

Figure 3. A conceptual framework for correlating the challenges in developing secure mHealth apps.



Based on the results of Table 2, we identified the most critical challenges for developing secure mHealth apps. Critical challenges can be determined if a specific challenge has a frequency of ≥50% of the selected studies. This criterion has been used by other researchers in different domains [72]. As in Table 2, the frequencies are shown for each challenge in the reviewed studies. By using this criterion, we concluded that there are 2 main critical challenges: lack of security guidelines for developing secure mHealth apps (20/32 studies, 63%) and developers' lack of security knowledge and expertise for secure mHealth app development (18/32 studies, 56%).

Despite the fact that other challenges were given less attention by the reviewed studies (ie, 19% for C3, 16% for C4, 13% for C5-C7, 9% for C8, and 6% for C9), some challenges have a direct relationship with other challenges as we indicated earlier (eg, poor security decisions during mHealth app development are related to insufficient security knowledge by developers). Consequently, there will be an impact on the development process of mHealth apps. Therefore, we believe identifying these challenges would help mHealth app development organizations evaluate their security practices and readiness in implementing security in mHealth app projects.

Discussion

Principal Findings

While mHealth apps enable health care services, the security of end users' health data remains a challenge. This review aimed to identify the challenges that prevent development of secure mHealth apps based on the existing literature. We identified 9 challenges based on the analysis of the data extracted from 32 articles. The identified challenges include (1) lack of security guidelines and regulations for developing secure mHealth apps, (2) developers' lack of knowledge of and expertise with secure mHealth app development, (3) lack of stakeholders' involvement during mHealth app development, (4) no or little attention by developers towards the security of mHealth apps, (5) lack of resources for developing secure mHealth apps, (6) project constraints during the mHealth app development process, (7) lack of security testing during mHealth app development, (8) developers' lack of motivation and ethical considerations, and (9) lack of security experts' engagement during mHealth app development. We noticed from the literature that there is an emphasis on presenting the security issues of mHealth apps and how they can be resolved (eg, presenting security framework, providing secure mHealth app development recommendations, evaluating the security for existing mHealth apps). However, little attention has been given to the human factor during the development process of mHealth apps (ie, nontechnical solutions). Hence, it would be critical to recognize the security

challenges that mHealth app developers face during the development process.

Sufficient security knowledge for mHealth app developers is one of the key factors that would help develop secure apps. Security knowledge can be discussed as the type of required security knowledge and the sources of acquiring that knowledge. According to Barnum and McGraw [52], there are 7 security knowledge categories for developing secure software, including knowledge of principles, guidelines, rules, attack pattern, vulnerability, exploit, and historical risk. While the presented set of security knowledge provides a strong foundation for enhancing security, it would be a bit challenging for developers since security knowledge is scattered all around. By considering security knowledge of vulnerabilities as an example, attackers can find a single vulnerability to exploit an app (ie, launching an attack). In contrast, developers should be aware of all security vulnerabilities and apply proper security measures and patches, which can be a daunting task. mHealth apps, more specifically, are connected to IoT devices, which makes securing the apps a challenge. Sikder et al [56] indicated that attackers could illegally access health data by exploiting sensors' permissions, which could enable them to extract data and transfer malware to an app. Therefore, further support for mHealth app developers' security knowledge is needed to cope with the rapid changes in security knowledge.

Likewise, using trusted sources (ie, tools and libraries) would be challenging for developers to be aware of their secure usage. So, we suggest further required improvement to facilitate mHealth app developers' jobs by exploring the list of trusted sources. Identifying trusted sources with their policies, terms, and conditions of usage and the proper ways of receiving updates would help mHealth app developers to develop secure apps. At the same time, this approach would help disseminate and provide security knowledge for mHealth app developers through trusted sources.

The Role of Security Experts Within mHealth App Development

"A critical challenge facing software security today is the dearth of experienced practitioners" [52]. A report by Ponemon Institute showed there is a dearth of security experts in mobile app development. Only 41% of the participants indicated that their organizations had sufficient security expertise [64]. Hence, having a security expert can be a strategic advantage for an organization. The role of security experts is quite crucial in developing secure mHealth apps. We conclude from the conceptual framework (ie, Figure 3) that a lack of security experts is already linked to most challenges. Without security experts on a team, the required security knowledge will be missing (ie, what security guidelines need to be followed, what security tools are available to be utilized, and which libraries can be trusted). As a result, developers' security knowledge would remain insufficient. Lack of security experts within mHealth app development organizations can lead to poor coding practices, rushing to deliver an app without even performing security testing. Furthermore, collaboration and social interactions with security experts and other team members would significantly impact security. As a result, removing the boundaries and stimulating common interests, in turn, support exchanging knowledge and ideas [67]. Also, it is good practice to exchange security knowledge, leverage that knowledge within the project, and acquire new knowledge.

Importance of Security Knowledge and Expertise to Develop Secure mHealth Apps

Our analysis shows that developers' lack of security knowledge and expertise for secure mHealth app development is correlated with most of the identified challenges. For instance, developing secure mHealth apps requires good knowledge about security guidelines, security tools, and the trusted libraries (ie, awareness of how, when, and why they should utilize them). It is worth mentioning that development of secure mHealth apps has become complex and challenging. mHealth apps require connection with external sensors or devices (eg, wearable devices, implantable devices) [56]. Nevertheless, providing the required learning resources can be underestimated by mHealth app development organizations [65]. Thus, organizations are required to provide security material to allow developers to learn to connect mHealth apps with emerging technologies (ie, IoT). Providing resources to support secure mHealth app development would contribute to filling the security knowledge gap and help open developers' mindset to security errors that need to be avoided [55].

Future Work

The results of our review enabled us to propose the following areas that warrant future research on the secure development of mHealth apps.

Challenges With and Practices of Developing Secure mHealth Apps With Real-World Practitioners

In this review, we identified the challenges that hinder developing secure mHealth apps based on SLR. We plan to conduct an empirical study to investigate the challenges with real-world practitioners to validate our results. The planned future research would enable us to compare the identified challenges identified from the literature with real-world practices for better understanding. Further, we aim to study the practices that real-world practitioners use to overcome the identified challenges. As a consequence, this would allow us to define which challenges are correlated with which practices. Hence, identifying the challenges and practices would help us to extend the current conceptual framework and provide a body of knowledge for secure mHealth app development.

Developers' Motivations and Ethical Considerations for Developing Secure mHealth Apps

Since motivations and ethical considerations play an essential role in the secure mHealth app development process, we assert that there is a need to conduct an empirical study to understand developers' motivational factors and what inspires them to ensure the security of mHealth apps (eg, security leaders, reward, recognition, career path, or promotion). Such a study can be further investigated by collecting quantitative data (eg, hypothesis testing) or qualitative data. This would create a better understanding and help mHealth app development organizations to realize and focus on the motivational factors.

Limitations

One of the potential threats for our SLR can be missing or excluding relevant studies. To mitigate this threat, we used Scopus library as our data source. Scopus is considered the largest indexing system that provides the most comprehensive search engine, among other digital libraries [73]. Scopus enabled us to get a reasonable number of studies (1867 articles). Furthermore, we tested our search string based on the pilot search to improve it and reach the relevant studies for this review. We selected the studies based on predefined inclusion and exclusion criteria. However, including and excluding studies can be impacted by researchers' subjective judgement. To mitigate this threat, the reasons for excluding the papers were recorded and reviewed by 2 independent researchers (who were previously mentioned).

Our research can be influenced by the researcher's bias in extracting data from the reviewed studies, which may negatively affect the findings. To overcome this threat, we extracted data based on a predefined data extraction form (see Multimedia Appendix 2). To mitigate the researcher's bias in data extraction and synthesis, the second author and the 2 independent researchers randomly verified the key points and themes derived by the first author through discussions.

Conclusion

This review was motivated by the growing amount of attention paid to mobile apps, particularly mHealth apps. We aimed to analyze and synthesize the literature to identify the challenges that hinder mHealth app developers from developing secure apps. Our review followed an SLR approach and selected 32 studies that we believed were relevant to our study. We identified and discussed 9 challenges faced by mHealth app developers to develop secure apps. We also provided a conceptual framework for the identified challenges and presented several challenges linked to the body of knowledge found in this literature review. Our findings can be valuable for researchers and practitioners (eg, mHealth app developers, managers) to support research and development of secure mHealth apps. For researchers, this review can help formulate and test hypotheses. Furthermore, ideal and innovative solutions can be proposed to address these challenges. For practitioners, our review can help understand the existing challenges for developing secure mHealth apps from the literature. This would help resolve these challenges at the early stages of the mHealth app development process.

Acknowledgments

We thank Dr Leonardo Horn Iwaya and Dr Faheem Ullah for playing the role of expert independent researchers during the key phases (eg, study selection, data extraction and synthesis) of this study. We also thank the anonymous reviewers for their valuable comments and suggestions that improved the paper. This work is partially funded by the Cyber Security Cooperative Research Centre.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of the reviewed studies.
[DOCX File , 18 KB - mhealth_v9i6e15654_app1.docx]

Multimedia Appendix 2
Data extraction form.
[DOCX File , 12 KB - mhealth_v9i6e15654_app2.docx]

References

1. Hussain M, Zaidan A, Zidan B, Iqbal S, Ahmed M, Albahri O, et al. Conceptual framework for the security of mobile health applications on Android platform. Telematics and Informatics 2018 Aug;35(5):1335-1354. [doi: 10.1016/j.tele.2018.03.005]
2. Müthing J, Jäschke T, Friedrich CM. Client-Focused Security Assessment of mHealth Apps and Recommended Practices to Prevent or Mitigate Transport Security Issues. JMIR Mhealth Uhealth 2017 Oct 18;5(10):e147 [FREE Full text] [doi: 10.2196/mhealth.7791] [Medline: 29046271]
3. Varshney U. Mobile health: Four emerging themes of research. Decision Support Systems 2014 Oct;66:20-35. [doi: 10.1016/j.dss.2014.06.001]
4. Aljedaani AA, Zahedi M, Babar MA. An Empirical Study on Developing Secure Mobile Health Apps: The Developers' Perspective. 2020 Presented at: 27th Asia-Pacific Software Engineering Conference (APSEC); December 1-4, 2020; Singapore. [doi: 10.1109/apsec51365.2020.00029]
5. Knorr K, Aspinall D. Security testing for Android mHealth apps. IEEE:pp 2015:1-8. [doi: 10.1109/icstw.2015.7107459]
6. 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. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]

7. Chin EM. Helping developers construct secure mobile applications. UC Berkeley 2013:10-22 PhD Thesis.
8. Research2guidance. https://research2guidance.com. 2017. URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2021-03-20]
9. Zubaydi F, Saleh A, Aloul F, Sagahyroon A. Security of mobile health (mHealth) systems. IEEE 2015:1-5. [doi: 10.1109/bibe.2015.7367689]
10. Weir C, Hermann B, Fahl S. From Needs to Actions to Secure Apps? The Effect of Requirements and Developer Practices on App Security. USENIX. 2020. URL: https://www.usenix.org/conference/usenixsecurity20/presentation/weir [accessed 2021-06-01]
11. Cifuentes Y, Beltrán L, Ramírez L. Analysis of security vulnerabilities for mobile health applications. International Journal of Health and Medical Engineering 2015;9(9):1067-1072. [doi: 10.5281/zenodo.1108987]
12. Adhikari R, Richards D, Scott K. Security and privacy issues related to the use of mobile health apps. 2014 Presented at: 25th Australasian Conference on Information Systems; December 8-10, 2014; Auckland, New Zealand URL: https://www.colleaga.org/sites/default/files/attachments/acis20140_submission_12.pdf
13. Flaten H, St Claire C, Schlager E, Dunnick C, Dellavalle R. Growth of mobile applications in dermatology - 2017 update. Dermatol Online J 2018 Feb 15;24(2):1 [FREE Full text] [Medline: 29630159]
14. Zahra F, Hussain A, Mohd H. Factor Affecting Mobile Health Application for Chronic Diseases. Journal of Telecommunication, Electronic and Computer Engineering 2018;10:77-81.
15. Mabo T, Swar B, Aghili S. A Vulnerability Study of Mhealth Chronic Disease Management (CDM) Applications (apps). In: Rocha Á, Adeli H, Reis LP, Costanzo S, editors. Trends and Advances in Information Systems and Technologies. WorldCIST'18 2018. Advances in Intelligent Systems and Computing, vol 745. Cham, Switzerland: Springer; 2018:587-598.
16. Ramey L, Osborne C, Kasitinon D, Juengst S. Apps and Mobile Health Technology in Rehabilitation: The Good, the Bad, and the Unknown. Phys Med Rehabil Clin N Am 2019 May;30(2):485-497. [doi: 10.1016/j.pmr.2018.12.001] [Medline: 30954161]
17. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
18. 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] [Medline: 25599627]
19. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and Security in Mobile Health: A Research Agenda. Computer (Long Beach Calif) 2016 Jun;49(6):22-30 [FREE Full text] [doi: 10.1109/MC.2016.185] [Medline: 28344359]
20. Papageorgiou A, Strigkos M, Politou E, Alepis E, Solanas A, Patsakis C. Security and Privacy Analysis of Mobile Health Applications: The Alarming State of Practice. IEEE Access 2018;6:9390-9403. [doi: 10.1109/ACCESS.2018.2799522]
21. Gejibo S, Mancini F, Mughal KA, Valvik RAB, Klungsøyr J. Secure data storage for mobile data collection systems. 2012 Presented at: International Conference on Management of Emergent Digital EcoSystems; October 2012; Addis Ababa Ethiopia p. 131-144. [doi: 10.1145/2457276.2457300]
22. Kanniah SL, Mahrin MN. A Review on Factors Influencing Implementation of Secure Software Development Practices. International Scholarly and Scientific Research & Innovation 2016;10(8):3032-3039 [FREE Full text] [doi: 10.5281/zenodo.1127256]
23. Thomas TW, Tabassum M, Chu B, Lipford H. Security During Application Development: an Application Security Expert Perspective. 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; April 2018; Montreal Quebec. [doi: 10.1145/3173574.3173836]
24. Raghavan V, Zhang X. Building security in during information systems development. In: AMCIS 2009 Proceedings. 2009 Presented at: Americas Conference on Information Systems; 2009; San Francisco, California p. 687-697.
25. Weir C, Rashid A, Noble J. Developer Essentials: Top Five Interventions to Support Secure Software Development. Lancaster University. 2017. URL: https://eprints.lancs.ac.uk/id/eprint/85933/4/InterventionsShortReport.pdf [accessed 2021-05-17]
26. Acar Y, Stransky C, Wermke D, Weir C, Mazurek ML, Fahl S. Developers Need Support, Too: A Survey of Security Advice for Software Developers. Developers Need Support, Too; 2017 Presented at: 2017 IEEE Cybersecurity Development (SecDev); September 24-26, 2017; Cambridge, MA. [doi: 10.1109/SecDev.2017.17]
27. Chatzipavlou IA, Christoforidou SA, Vlachopoulou M. A recommended guideline for the development of mHealth Apps. Mhealth 2016 May 18;2:21-21 [FREE Full text] [doi: 10.21037/mhealth.2016.05.01] [Medline: 28293597]
28. Katusiime J, Pinkwart N. A review of privacy and usability issues in mobile health systems: Role of external factors. Health Informatics J 2019 Sep 09;25(3):935-950 [FREE Full text] [doi: 10.1177/1460458217733121] [Medline: 28992749]
29. Marquez G, Astudillo H, Taramasco C. Security in Telehealth Systems From a Software Engineering Viewpoint: A Systematic Mapping Study. IEEE Access 2020;8:10933-10950. [doi: 10.1109/access.2020.2964988]
30. Kitchenham B, Pretorius R, Budgen D, Pearl Brereton O, Turner M, Niazi M, et al. Systematic literature reviews in software engineering – A tertiary study. Information and Software Technology 2010 Aug;52(8):792-805. [doi: 10.1016/j.infsof.2010.03.006]

31. Cruzes D, Dyba T. Recommended steps for thematic synthesis in software engineering. IEEE 2011:275-284. [doi: 10.1109/esem.2011.36]
32. He D, Naveed M, Gunter CA, Nahrstedt K. Security Concerns in Android mHealth Apps. AMIA Annu Symp Proc 2014;2014:645-654 [FREE Full text] [Medline: 25954370]
33. Regoniel P. Conceptual framework: A step by step guide on how to make one. SimplyEducate Me 2015:NA-NA.
34. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht UV. mHealth 2.0: Experiences, Possibilities, and Perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
35. Mergel I. The Long Way From Government Open Data to Mobile Health Apps: Overcoming Institutional Barriers in the US Federal Government. JMIR Mhealth Uhealth 2014 Dec 23;2(4):e58 [FREE Full text] [doi: 10.2196/mhealth.3694] [Medline: 25537314]
36. Carter A, Liddle J, Hall W, Chenery H. Mobile Phones in Research and Treatment: Ethical Guidelines and Future Directions. JMIR Mhealth Uhealth 2015 Oct 16;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
37. Landman A, Emani S, Carlile N, Rosenthal DI, Semakov S, Pallin DJ, et al. A mobile app for securely capturing and transferring clinical images to the electronic health record: description and preliminary usability study. JMIR Mhealth Uhealth 2015 Jan 02;3(1):e1 [FREE Full text] [doi: 10.2196/mhealth.3481] [Medline: 25565678]
38. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The Top Chinese Mobile Health Apps: A Systematic Investigation. J Med Internet Res 2016 Aug 29;18(8):e222 [FREE Full text] [doi: 10.2196/jmir.5955] [Medline: 27573724]
39. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth Assessment: Conceptualization of a Global Framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
40. Al-Sharo YM. Networking Issues for Security and Privacy in Mobile Health Apps. IJACSA 2019;10(2):1. [doi: 10.14569/IJACSA.2019.0100225]
41. 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] [Medline: 31045497]
42. Muchagata J, Teles S, Vieira-Marques P, Abrantes D, Ferreira A. Dementia and mHealth: On the Way to GDPR Compliance. In: Roque A, Tomczyk A, De Maria E, Putze F, Moucek R, Fred A, et al, editors. Biomedical Engineering Systems and Technologies. Switzerland: Springer International Publishing; 2020.
43. Luxton DD, Kayl RA, Mishkind MC. mHealth data security: the need for HIPAA-compliant standardization. Telemed J E Health 2012 May;18(4):284-288. [doi: 10.1089/tmj.2011.0180] [Medline: 22400974]
44. Aliasgari M, Black M, Yadav N. Security Vulnerabilities in Mobile Health Applications. 2019 Presented at: IEEE Conference on Application, Information and Network Security (AINS); November 21-22, 2018; Langkawi, Malaysia.
45. Zens M, Südkamp NP, Niemeyer P. „Back on Track“: Kreuzbandstudie via Smartphone. Arthroskopie 2016 Oct 19;29(4):225-231. [doi: 10.1007/s00142-016-0098-6]
46. Thamilarasu G, Lakin C. A Security Framework for Mobile Health Applications. 2017 Presented at: 5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW); August 21-23, 2017; Prague, Czech Republic URL: https://ieeexplore.ieee.org/document/8113797
47. Parker L, Halter V, Karliychuk T, Grundy Q. How private is your mental health app data? An empirical study of mental health app privacy policies and practices. Int J Law Psychiatry 2019;64:198-204. [doi: 10.1016/j.ijlp.2019.04.002] [Medline: 31122630]
48. Srivastava M, Thamilarasu G. MSF: A Comprehensive Security Framework for mHealth Applications. 2020 Presented at: 7th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW); August 26-28, 2019; Istanbul, Turkey URL: https://ieeexplore.ieee.org/abstract/document/9052155
49. Hutton L, Price BA, Kelly R, McCormick C, Bandara AK, Hatzakis T, et al. Assessing the Privacy of mHealth Apps for Self-Tracking: Heuristic Evaluation Approach. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e185 [FREE Full text] [doi: 10.2196/mhealth.9217] [Medline: 30348623]
50. Arora S, Yttri J, Nilse W. Privacy and Security in Mobile Health (mHealth) Research. Alcohol Res 2014;36(1):143-151 [FREE Full text] [Medline: 26259009]
51. Wykes T, Lipshitz J, Schueller SM. Towards the Design of Ethical Standards Related to Digital Mental Health and all Its Applications. Curr Treat Options Psych 2019 Jul 5;6(3):232-242. [doi: 10.1007/s40501-019-00180-0]
52. Barnum S, McGraw G. Knowledge for Software Security. IEEE Secur. Privacy Mag 2005 Mar;3(2):74-78. [doi: 10.1109/msp.2005.45]
53. Nurgalieva L, O'Callaghan D, Doherty G. Security and Privacy of mHealth Applications: A Scoping Review. IEEE Access 2020;8:104247-104268. [doi: 10.1109/access.2020.2999934]
54. Assal H, Chiasson S. 'Think secure from the beginning' A Survey with Software Developers. In: Proceedings of the CHI conference on human factors in computing systems. 2019 Presented at: CHI conference on human factors in computing systems; 2019; UK p. 1-13. [doi: 10.1145/3290605.3300519]

55. Weir C, Rashid A, Noble J. How to improve the security skills of mobile app developers? Comparing and contrasting expert views. In: Symposium on Usable Privacy and Security. 2016 Presented at: In Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016); 2016; Denver, Colorado.
56. Sikder A, Petracca G, Aksu H, Jaeger T, Uluagac A. A Survey on Sensor-based Threats to Internet-of-Things (IoT) Devices and Applications. arXiv preprint arXiv 2018:1802.
57. Faily S. Engaging Stakeholders in Security Design: An Assumption-Driven Approach. In: Proceedings of the 8th International Symposium on Human Aspects of Information Security & Assurance. 2014 Presented at: the 8th International Symposium on Human Aspects of Information Security & Assurance; 2014; Plymouth.
58. Wurster G, van Oorschot PC. The developer is the enemy. In: ACM. 2009 Presented at: In Proceedings of the 2008 New Security Paradigms Workshop; 2008; Lake Tahoe, CA, USA p. 89-97. [doi: 10.1145/1595676.1595691]
59. Nguyen D, Wermke D, Acar Y, Backes M, Weir C, Fahl S. A Stitch in Time: Supporting Android Developers in Writing Secure Code. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017 Presented at: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017; Dallas, TX, USA p. 1065-1077. [doi: 10.1145/3133956.3133977]
60. Balebako R, Marsh A, Lin J, Hong J, Cranor L. The privacy and security behaviors of smartphone app developers. In: Proceedings of Workshop on Usable Security (USEC). 2014 Presented at: Proceedings of Workshop on Usable Security (USEC); 2014; San Diego, California. [doi: 10.14722/usec.2014.23006]
61. McGraw G. Software Security: Building Security In. In: Book. Boston, MI, USA: Addison-Wesley Professional; Jan 23, 2006.
62. Lipner S. Security development lifecycle. In: Book. USA: Microsoft Press; Jun 24, 2006:135-137.
63. Woon IM, Kankanhalli A. Investigation of IS professionals' intention to practise secure development of applications. International Journal of Human-Computer Studies 2007 Jan;65(1):29-41. [doi: 10.1016/j.ijhcs.2006.08.003]
64. Ponemon IL. The State of Mobile Application Insecurity. Website. URL: https://www.workplaceprivacyreport.com/wp-content/uploads/sites/162/2015/03/WGL03074USEN.pdf [accessed 2021-05-20]
65. National CSC. Secure development is everyone's concern. Website. 2017. URL: https://www.ncsc.gov.uk/guidance/secure-development-everyones-concern [accessed 2021-05-20]
66. N Inukollu V, Keshamon DD, Kang T, Inukollu M. Factors Influncing Quality of Mobile Apps: Role of Mobile App Development Life Cycle. IJSEA 2014 Sep 30;5(5):15-34. [doi: 10.5121/ijsea.2014.5502]
67. Sharp H, Baddoo N, Beecham S, Hall T, Robinson H. Models of motivation in software engineering. Information and Software Technology 2009 Jan;51(1):219-233. [doi: 10.1016/j.infsof.2008.05.009]
68. Beautement A, Sasse M, Wonham M. The compliance budget: managing security behaviour in organisations. In: Proceedings of the 2008 New Security Paradigms Workshop.: ACM; 2009 Presented at: Proceedings of the 2008 New Security Paradigms Workshop; 2008; Lake Tahoe, CA, USA p. 47-58.
69. Verner J, Babar M, Cerpa N, Hall T, Beecham S. Factors that motivate software engineering teams: A four country empirical study. Journal of Systems and Software 2014 Jun;92:115-127. [doi: 10.1016/j.jss.2014.01.008]
70. Xie J, Lipford H, Chu B. Why do programmers make security errors? 2011 Presented at: 2011 IEEE symposium on visual languages and human-centric computing (VL/HCC); 2011; Pittsburgh, PA p. 161-164. [doi: 10.1109/vlhcc.2011.6070393]
71. Jabareen Y. Building a Conceptual Framework: Philosophy, Definitions, and Procedure. International Journal of Qualitative Methods 2009 Dec;8(4):49-62. [doi: 10.1177/160940690900800406]
72. Khan AA, Keung J, Hussain S, Niazi M, Tamimy MMI. Understanding software process improvement in global software development. SIGAPP Appl. Comput. Rev 2017 Aug 03;17(2):5-15. [doi: 10.1145/3131080.3131081]
73. Shahin M, Ali Babar M, Zhu L. Continuous Integration, Delivery and Deployment: A Systematic Review on Approaches, Tools, Challenges and Practices. IEEE Access 2017;5:3909-3943. [doi: 10.1109/access.2017.2685629]

Abbreviations

GDPR: General Data Protection Regulation
HIPAA: Health Insurance Portability and Accountability Act
IoT: Internet of Things
mHealth: mobile health
OWASP: Open Web Application Security Project
SDLC: software development lifecycle
SLR: systematic literature review

Edited by L Buis; submitted 26.07.19; peer-reviewed by L Zhou, L Parker, M Diefenbach, L Fleisher; comments to author 25.08.20; revised version received 13.01.21; accepted 16.04.21; published 21.06.21.

Please cite as:
Aljedaani B, Babar MA
Challenges With Developing Secure Mobile Health Applications: Systematic Review
JMIR Mhealth Uhealth 2021;9(6):e15654
URL: https://mhealth.jmir.org/2021/6/e15654
doi:10.2196/15654
PMID:34152277

Original Paper

Quality of Physical Activity Apps: Systematic Search in App Stores and Content Analysis

Sarah Paganini[1], PhD; Yannik Terhorst[2], MSc; Lasse Bosse Sander[3], PhD; Selma Catic[2], BSc; Sümeyye Balci[2], MSc; Ann-Marie Küchler[2], MSc; Dana Schultchen[4], PhD; Katrin Plaumann[5], MSc; Sarah Sturmbauer[6], MSc; Lena Violetta Krämer[3], PhD; Jiaxi Lin[7], PhD; Ramona Wurst[1], MSc; Rüdiger Pryss[8], Prof Dr, PhD; Harald Baumeister[2], Prof Dr, PhD; Eva-Maria Messner[2], PhD

[1]Department of Sport Psychology, Institute of Sports and Sport Science, Albert-Ludwigs-University Freiburg, Freiburg, Germany

[2]Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Ulm, Germany

[3]Department of Rehabilitation Psychology and Psychotherapy, Institute of Psychology, Albert-Ludwigs-University Freiburg, Freiburg, Germany

[4]Department of Clinical and Health Psychology, Institute of Psychology and Education, Ulm University, Ulm, Germany

[5]Institute of Media Informatics, Ulm University, Ulm, Germany

[6]Chair of Health Psychology, Department of Psychology, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

[7]Department of Psychiatry and Psychotherapy Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[8]Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany

Corresponding Author:
Eva-Maria Messner, PhD
Department of Clinical Psychology and Psychotherapy
Institute of Psychology and Education
Ulm University
Lise-Meitner-Str 16
Ulm, 89081
Germany
Phone: 49 7315032802
Email: eva-maria.messner@uni-ulm.de

Abstract

Background: Physical inactivity is a major contributor to the development and persistence of chronic diseases. Mobile health apps that foster physical activity have the potential to assist in behavior change. However, the quality of the mobile health apps available in app stores is hard to assess for making informed decisions by end users and health care providers.

Objective: This study aimed at systematically reviewing and analyzing the content and quality of physical activity apps available in the 2 major app stores (Google Play and App Store) by using the German version of the Mobile App Rating Scale (MARS-G). Moreover, the privacy and security measures were assessed.

Methods: A web crawler was used to systematically search for apps promoting physical activity in the Google Play store and App Store. Two independent raters used the MARS-G to assess app quality. Further, app characteristics, content and functions, and privacy and security measures were assessed. The correlation between user star ratings and MARS was calculated. Exploratory regression analysis was conducted to determine relevant predictors for the overall quality of physical activity apps.

Results: Of the 2231 identified apps, 312 met the inclusion criteria. The results indicated that the overall quality was moderate (mean 3.60 [SD 0.59], range 1-4.75). The scores of the subscales, that is, *information* (mean 3.24 [SD 0.56], range 1.17-4.4), *engagement* (mean 3.19 [SD 0.82], range 1.2-5), a*esthetics* (mean 3.65 [SD 0.79], range 1-5), and *functionality* (mean 4.35 [SD 0.58], range 1.88-5) were obtained. An efficacy study could not be identified for any of the included apps. The features of data security and privacy were mainly not applied. Average user ratings showed significant small correlations with the MARS ratings (r=0.22, 95% CI 0.08-0.35; P<.001). The amount of content and number of functions were predictive of the overall quality of these physical activity apps, whereas app store and price were not.

Conclusions: Apps for physical activity showed a broad range of quality ratings, with moderate overall quality ratings. Given the present privacy, security, and evidence concerns inherent to most rated apps, their medical use is questionable. There is a need for open-source databases of expert quality ratings to foster informed health care decisions by users and health care providers.

(JMIR Mhealth Uhealth 2021;9(6):e22587) doi:10.2196/22587

KEYWORDS

sports; exercise; mobile apps; mHealth; quality indicators; systematic review

Introduction

Physical inactivity is a significant risk factor for noncommunicable diseases such as cancer, diabetes, cardiovascular diseases, or chronic respiratory diseases and is estimated to cause 6%-10% of these diseases worldwide [1,2]. Insufficient physical activity is also a leading risk factor for mortality and was reported to be associated with 9% of premature death cases in 2008 [2]. The World Health Organization recommends at least 150 minutes of moderate or 75 minutes of vigorous-intensity physical activity per week for adults [3]. However, about 30% of adults do not follow this recommendation and are physically inactive [4].

Evidence indicates that regular physical activity results in physical, social, and mental health benefits such as better quality of sleep, lower depressive symptomatology, higher well-being, and a reduced risk of a large number of noncommunicable diseases [5-7]. Mobile apps might be a cost-effective and scalable option to foster behavior change in daily life [8]. Apps can also be beneficial as a supplement to behavioral interventions [9]. Additionally, fitness apps are very popular in the general population; a survey conducted in the United States in 2015 showed that about 58% of mobile phone users had downloaded a health app [10]. Of these, the most common categories were fitness and nutrition apps, and most respondents were using them daily. Moreover, health app users were more likely to meet the World Health Organization recommendations concerning physical activity [3,11] and apps were found to be efficacious in promoting physical activity with moderate effect sizes [12]. In 2 recent meta-analyses of apps for increasing physical activity, there was an increase in objectively measured physical activity in the app groups compared to that in the control groups [13,14]. However, these differences were not significant. Regarding the content and quality of apps promoting physical activity, previous reviews focused mainly on the use of behavioral change techniques (BCTs) developed by Abraham and Michie [15-21]. The most often provided BCTs were feedback on performance, self-monitoring, and goal setting [15,16,18,19].

Regarding the quality of apps, Schoeppe and colleagues [22] used the standardized Mobile App Rating Scale (MARS) [23] to evaluate apps for improving diet, physical activity, and sedentary behavior. However, the mentioned reviews show some limitations as these reviews mostly evaluated apps with specific characteristics (eg, only apps that are connected to an electronic activity monitor) [18], apps especially developed for children and adolescents [20,22], a limited number of apps (eg, only the 20 top-ranked apps, random selection of apps, or apps with a star-rating of at least 4) [16,19,22], or apps with certain contents and features (eg, only apps with feedback and apps that follow the official World Health Organization recommendation for physical activity) [15]. Overall, most studies evaluating apps for health behavior change use self-developed evaluation checklists and do not assess privacy and security features [24]. The only validated evaluation tools that were used are the BCT taxonomy and, in a few studies, the MARS [15,16,18,22,24].

Besides the evaluation of theory-based content, applied techniques or functions, and effectiveness, it is important to consider the risks of mobile health app use, such as inadequate protection of data and privacy or lack of informed consent [24,25]. Many physical activity apps are available, particularly via the 2 largest app stores, Google Play store and the App Store but there is only limited information on the quality and data security of these apps [24,26]. Initial studies concerning data security of medicine-related, depression, and smoking cessation apps reveal worrying results as sharing medical health data with third parties is routine and mostly not made transparent [27,28]. The only study evaluating the safety of personal data in physical activity apps also revealed substantial shortcomings [19]. In terms of quality, user ratings seem to be a questionable indicator as they seem to be mostly influenced by usability and functionality [26,29]. However, a recent evaluation revealed a positive correlation between a broad range of app quality ratings and user star ratings [30].

The aim of this study was to conduct a systematic and objective investigation of the physical activity apps available in 2 major app stores by using the German version of MARS (MARS-G). The MARS-G is a multidimensional instrument specifically developed to assess app quality on the dimensions engagement, functionality, aesthetics, and information quality [31]. Furthermore, privacy and security measures as well as the general characteristics and functions of physical activity apps will be assessed. The following research questions are addressed:

1. What is the quality of the apps promoting physical activity regarding engagement, functionality, aesthetics, and information?
2. What are the general characteristics, content, functions, privacy, and security measures of the apps promoting physical activity?
3. Are the user ratings in agreement with the expert quality ratings?
4. Which app features can predict app quality?

Methods

Search Strategy and Eligibility Criteria

A web crawler (automated web search engine) was used to scan the European Google Play store and App Store to search for eligible apps. The search was carried out on February 20, 2018 by using the following search terms: (1) active, (2) endurance, (3) exercise, (4) fitness, (5) gymnastics, (6) muscle, (7) shape, (8) strength, (9) training, and (10) workout. The search string to identify apps targeting physical activity was developed in an expert discussion (EMM and HB). The web crawler searches for each term and app store. Duplicates were automatically

removed. After identification, a two-step procedure was applied by 2 independent researchers: (1) checking eligibility based on app title and description and (2) checking eligibility based on information in the downloaded app. In the first step, all identified apps were screened for whether their title, description, and images indicated that the app was developed for promoting physical activity (with at least 50% of the content focusing on physical activity); the app was available in German or English; the app was downloadable through the official Google Play store or App Store; the app could be used without further equipment, devices, or programs; and the app was primarily developed for adults. In the second step, all downloaded apps were assessed in detail to check whether they met the abovementioned eligibility criteria. If apps did not work after the download (checked with 2 different mobile phones) or were explicitly developed for children (explicitly stated in the title, description, or aims of the app), they were excluded. The other exclusion criteria were (1) app bundle, (2) only working with additional device (eg, Garmin connect), or (3) targeting specific person groups (eg, employees of a specific company).

Rating Procedure

Each app was rated by 2 independent raters between February and October 2018. All raters undertook a free online training [32] (training module last updated on November 25, 2019). Raters were recruited from an interdisciplinary expert team (sports science, sport psychology, clinical psychology, information technology: EMM, YT, SP, LS, JL, SC, SB, LK, AK, DS, SS, KP, RP, and RW). Each app was tested and used for at least 15-20 minutes before the rating. The interrater reliability between the raters was computed for quality assurance.

Outcome Measures

The MARS includes a multidimensional quality rating consisting of 4 dimensions: engagement (5 items: fun, interest, individual adaptability, interactivity, and target group), functionality (4 items: performance, usability, navigation, and gestural design), aesthetics (3 items: layout, graphics, and visual appeal), and information quality (7 items: accuracy of app description, goals, quality of information, quantity of information, quality of visual information, credibility, and evidence base) [23,31]. Hereby, the evidence base (dimension: information quality) was identified by app description and developer's or provider's websites. Items are rated from 1 (inadequate) to 5 (excellent). Besides these objective scales, subjective quality (recommendation, frequency of use, willingness to pay, overall star rating) and perceived impact on the user (awareness, knowledge, attitudes, intention to change, help-seeking, behavioral change) were assessed. Furthermore, the assessment includes a classification section to examine the app characteristics. The following variables were extracted: (1) app name, (2) store link, (3) platform (Google Play store and App Store), (4) content-related subcategory, (5) aims, (6) price, (7) user rating, (8) content, strategies, and functions (abbreviated as *functions* in the following, assessed with 22 items; ie, information/education, monitoring/tracking, goal setting, gamification, reminder) and (9) privacy and security features [23,31]. Privacy and security features were rated on a descriptive level (ie, presence of privacy policy, contact information or imprint, log-in with a password). Only information that was displayed within the app was used for evaluation. In this study, MARS-G was used [23,31]. The validation of the MARS-G yielded excellent internal consistency (ω=0.84, 95% CI 0.77-0.88) and high levels of interrater reliability (intraclass correlation coefficient [ICC] 0.83, 95% CI 0.82-0.85) [31].

Statistical Analysis

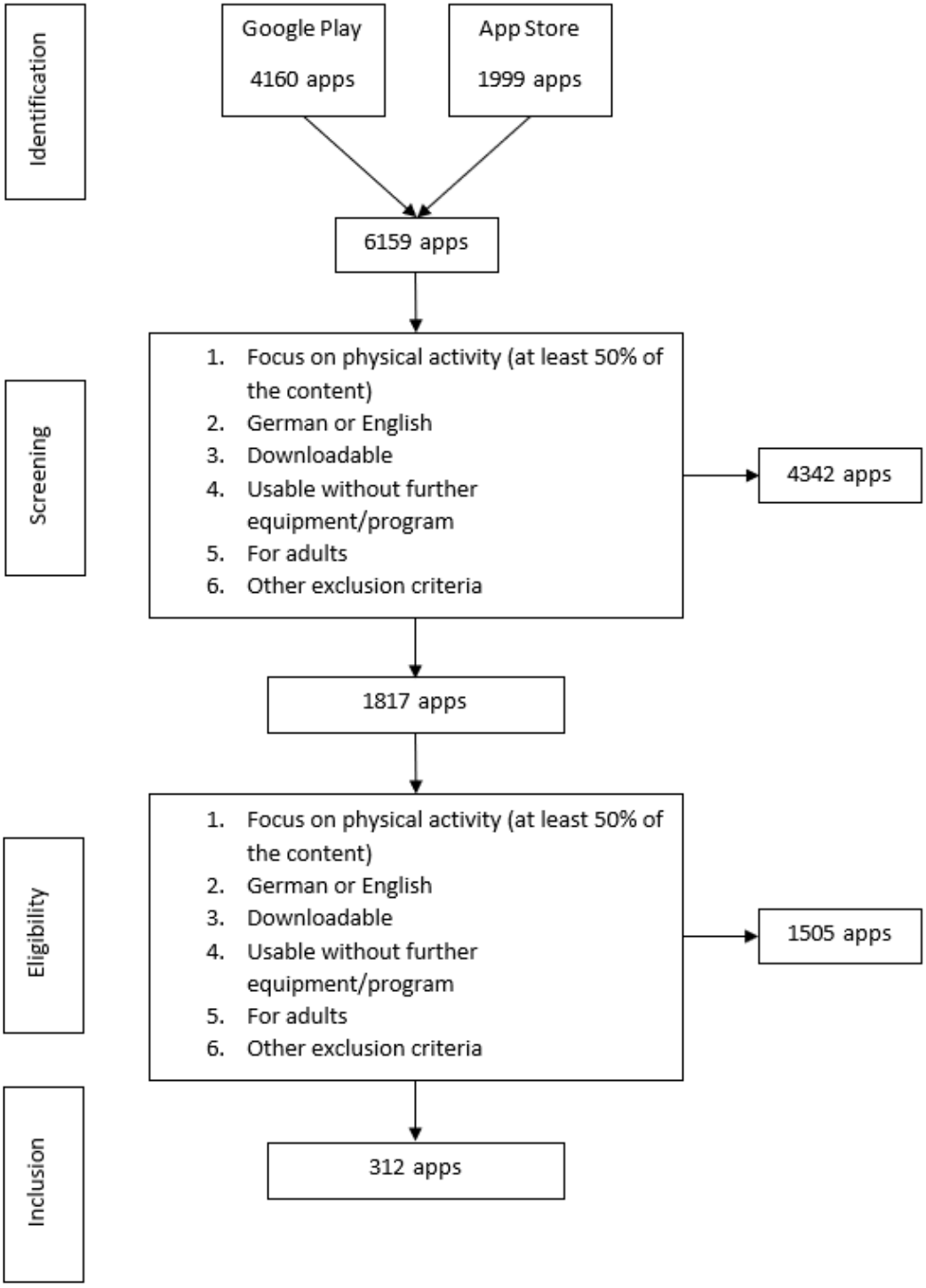
To evaluate consistency, the ICC between the raters was calculated for quality assurance. Rater agreement was examined by ICC based on a two-way mixed-effects model [33]. An ICC of <0.50 is considered poor, 0.51-0.75 as moderate, 0.76-0.89 as good, and >0.90 as excellent [34]. A minimum ICC of 0.8 was predefined as a sufficient ICC in this study. For quality evaluation, means and standard deviations were calculated for each dimension of the MARS separately and overall. For all calculations, the mean of both raters was used. Further correlation between user ratings provided by Google Play/App Store and the MARS rating was calculated. For correlations analysis, an alpha level of 5% was defined. *P* values were adjusted using the procedure proposed by Holm [35]. To determine relevant predictors for overall quality, exploratory multiple linear regression analysis was conducted. Price, store, and the number of functions were used as predictors, as they were significant predictors in other systematic app reviews (eg, older adults, mindfulness, depression, posttraumatic stress disorder, rheumatoid arthritis) [36-40]. Dichotomous predictors were dummy coded. Regression estimates represent unstandardized regression coefficients.

Results

Search Results

The search in the Google Play store and App Store yielded 6159 apps without duplicates. Screening resulted in the inclusion of 1817 apps. After downloading and assessing the eligibility criteria in detail, 1495 apps had to be excluded. The remaining 312 apps were included in the analyses (see Figure 1 for further information).

Figure 1. Flowchart of the inclusion and exclusion processes.

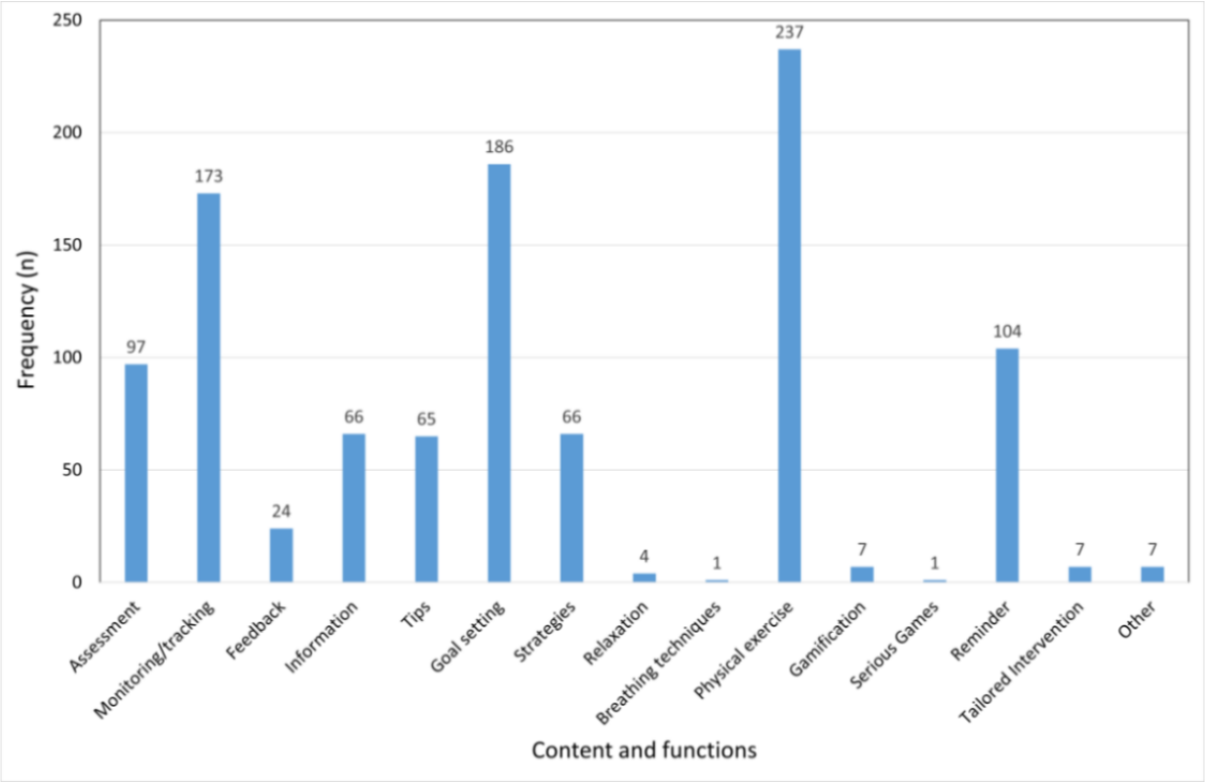


App Characteristics

Of the 312 included apps, 143 (45.8%) were available in Google Play store and 169 (54.2%) were available in the App Store. The average user's star rating was 4.29 (SD 0.70) (range 0-5). The costs for the included apps ranged from free for 96.8% (302/312) of the apps to €10.99 (mean cost €3.54 [SD 2.97]; €1=US $1.20). Regarding the store categories, the majority of the apps (300/312) were listed under health and fitness (multiple categories were assigned in stores). The further assigned categories were sports (n=83), lifestyle (n=69), business (n=4), entertainment (n=3), medicine (n=2), social (n=2), food and drink (n=1), and travel (n=1). Apps provided a broad range of functions (Figure 2). Most apps offered physical exercises (237/312, 75.9%) followed by goal setting (186/312, 59.6%) and monitoring/tracking (173/312, 55.4%). Apps further provided reminders (104/312, 33.3%), assessments (97/312, 31.1%), information (66/312, 21.2%), strategies/skills (66/312,

21.2%), tips/advice (65/312, 20.8%), feedback (24/312, 7.7%), gamification (7/312, 2.2%), tailoring (7/312, 2.2%), relaxation (4/312, 1.3%), breathing techniques (1/312, 0.3%), serious games (1/312, 0.3%), and others (7/312, 2.2%). On average, an app provided 3.34 (SD 2.29) functions (range 0-10). App description and developers' or providers' websites indicated that no app was certified according to the medical device law.

Figure 2. Content and functions of the included apps.



Data Security and Privacy Features

Of the 312 assessed apps, 67 (21.5%) had an imprint/contact information, 60 (19.2%) provided a visible privacy policy, 25 (8.0%) were only accessible with a personalized log-in, 20 (6.4%) utilized a passive informed consent (eg, by continuing you accept our privacy policy), 15 (4.8%) contained an active informed consent (eg, active opt-in to data collection/transfer), 16 (4.8%) had a password option, and 5 (1.6%) gave information about the financial background or made conflicts of interest transparent. No app had embedded emergency features (eg, in case of an accident).

App Quality

The ICC agreement between the raters was high (ICC 0.87, 95% CI 0.86-0.874). The average overall quality of the apps for physical activity was 3.60 (SD 0.59) (range 1-4.75) with average quality of information (3.24 [SD 0.56], range 1.17-4.4) and engagement (3.19 [SD 0.82], range 1.2-5). Aesthetics (3.65 [SD 0.79], range 1-5) was good and functionality was excellent (4.35 [SD 0.58], range 1.88-5). Of all the 312 included apps, 10 (3.2%) reached a MARS score of above 4.5. The evidence item (based on app description, developer's and provider's websites) indicated that no app was scientifically evaluated. The MARS quality ratings are summarized in Figure 3. The overall subjective quality reached an average rating of 2.34 (SD 0.78) and the overall perceived impact on the user was rated as 2.32 (SD 0.60). The details can be found in Table 1. The 10 apps with the highest quality ratings are presented in Multimedia Appendix 1.

Figure 3. Quality of the apps. MARS: Mobile App Rating Scale.

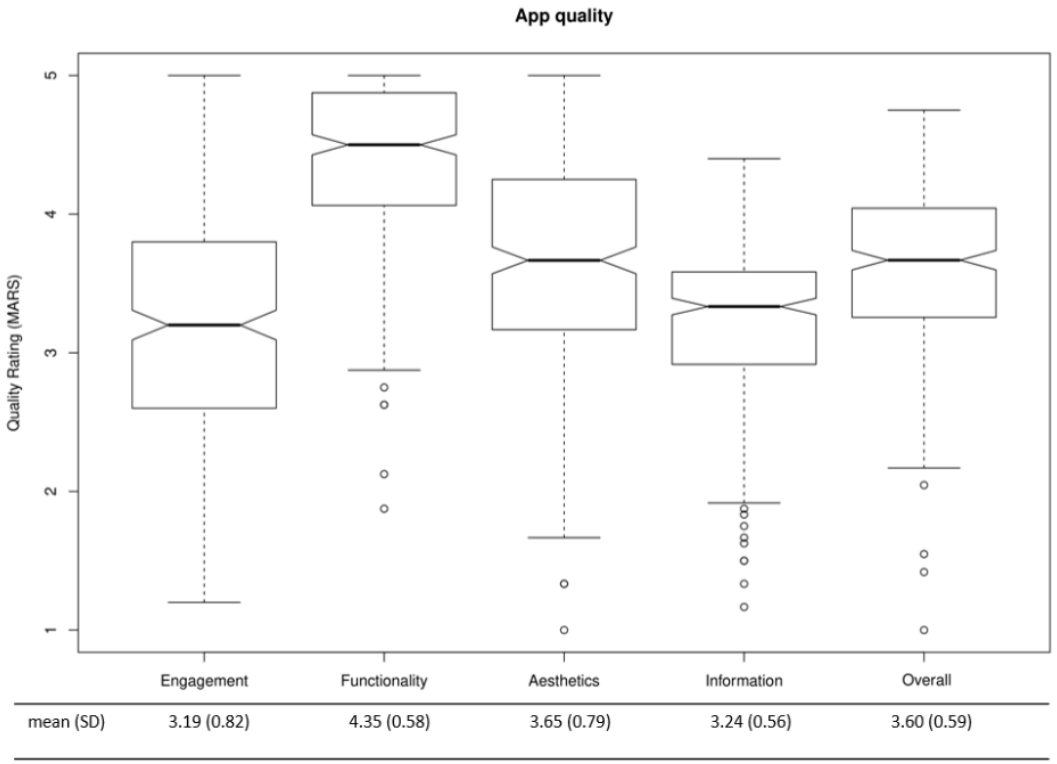


Table 1. Subjective quality ratings and ratings of the perceived impact on the users of mobile app rating scale.

Variable	Score, mean (SD)
Subjective quality rating	
Recommendable[a]	2.67 (1.05)
Use next 12 months[b]	2.38 (1.12)
Pay[c]	1.27 (0.46)
Star rating[d]	3.06 (0.89)
Overall	2.34 (0.78)
Perceived impact on user	
Increase awareness	2.16 (0.81)
Increase knowledge	2.65 (0.90)
Attitudes	2.26 (0.84)
Fosters intention to change	2.47 (0.90)
Empowers help-seeking	1.31 (0.62)
Fosters behavior change	3.07 (0.88)
Overall	2.32 (0.60)

[a]Rated on a 5-point scale (from 1: I would not recommend the app to anyone, to 5: I would recommend the app to everyone).

[b]Rated on a 5-point scale (1: never; 2: 1-2; 3: 3-10; 4: 10-50; 5: >50).

[c]Rated on a 3-point scale (1: I would not buy this app if it cost anything; 2: I might buy this app if it cost anything; 3: I would buy this app if it cost anything).

[d]Rated on a 5-point scale (from 1: one of the worst apps I have ever used, to 5: one of the best apps I have ever used).

User and Expert Agreement Toward Quality

Small correlations were found between the user ratings in the stores and the MARS. The correlations are summarized in Table 2.

Table 2. Correlation between mobile app rating scale and user rating.

Mobile app rating scale dimension	User rating, *r* (95% CI)	*P* value[a]
Engagement	0.25 (0.11-0.38)	<.001
Functionality	0.15 (0.01-0.28)	.03
Aesthetics	0.14 (0.02-0.24)	.03
Information quality	0.14 (0.02-0.26)	.03
Overall	0.22 (0.08-0.35)	<.001

[a]Adjusted *P* value for multiple testing [35].

Exploratory Regression Analysis

Exploratory regression analysis indicated that app quality could be predicted by the number of functions integrated into the app. Price and store had no predictive value. The results of the regression analyses are summarized in Table 3.

Table 3. Results of the regression analysis showing the predictors of app quality.

Predictor	β	SD	*t (df)*	*P* value	Adjusted R^2 (%)
Price	.22	0.19	1.15 (310)	.25	0.00
Store	–.13	0.07	–1.94 (310)	.05	0.88
Number of functions/amount of content	.12	0.01	9.64 (310)	<.001	22.82

Discussion

Principal Findings

In this study, the quality, general characteristics, privacy and security features, and content/functions of apps that promote physical activity in the commercial European app stores were systematically assessed. The included 312 apps showed a moderate overall quality (3.60 [SD 0.59], range 1-4.75). Moreover, several apps showed very high ratings, and there was a large range of quality ratings. The assessments of the 10 best-rated apps are described in detail in Multimedia Appendix 1. Functionality was the dimension with the highest rating, followed by aesthetics, information quality, and engagement. These results corroborate those of Schoeppe and colleagues [22] who evaluated diet and physical activity apps for children and adolescents (overall quality, mean 3.6).

The apps offered a variety of different functions (15 out of 22 functions were used). On average, 3 functions were applied per app, and the most common ones were exercises, goal setting, and monitoring/tracking. This is partly in line with previous reviews for apps promoting physical activity [15,16,18] or weight management [29]. Functions differed from the most frequently used BCTs in another review that evaluated apps for diet, physical activity, and sedentary behavior developed for children and adolescents (the top 3 functions provided instructions, general encouragement, contingent [22]). This discrepancy might be explained by the different target groups. Studies have already shown that BCTs incorporated in apps for health behavior change differ between adults and children/adolescents [20]. The average number of the 3 applied functions was lower compared to that in other reviews that reported 5-8 comparable BCTs [15,16,19,22]. This could be due to the broader range of the included apps in this study.

No randomized controlled trial evaluating the effectiveness of one of the included apps could be identified. This lack of a solid evidence base for the use of health apps is in line with that reported in other systematic reviews of app quality (eg, older adults, mindfulness, depression, rheumatoid arthritis, and posttraumatic stress disorder) [36-40]. These systematic reviews showed that the proportion of the scientifically evaluated apps ranges between 0% and 4.8%. Overall, this indicates a gap between research and health practices. Although there are several randomized controlled trials that investigate the efficacy of sport app use to foster behavior change, these apps are not available in the app stores [41,42]. Of note, a vast majority of apps are downloaded from the Google Play store and App Store [43]. This might stem from the lack of sustainable structures at universities (eg, end of funding, frequent job changes). Furthermore, this imposes a risk for safe sport app use as the evidence base is the gold standard for assuring quality and efficacy. Moreover, data privacy and security features were also rated as low. Only 19.2% (60/312) of all the apps provided a privacy policy, and 21.5% (67/312) of the included apps provided any contact information or an imprint. All other privacy and security features were fulfilled by less than 20% (range 0-60) of all the apps. In contrast, Bondaronek and colleagues [19] stated that almost 70% of the 65 included physical activity apps had a privacy policy. This might be because they searched for the best-ranked apps. Taken together, the ratings of information quality (including correctness, credibility, and scientific evidence) and the ratings of data

protection (including privacy policy, imprint, log-in, informed consent, password, conflicts of interest) reveal potential risks such as misinformation, adverse effects of app use, data misuse, or potential nonefficacy.

Average user ratings in the stores showed a significantly small correlation with the MARS ratings, which is in line with that reported in previous research [30,36]. However, several studies (including apps for weight management and chronic pain) could not identify an association [29,44]. This indicates that although user star ratings of physical activity apps might be used as an indicator for app quality, such an association should be evaluated for each indication separately. The results of our study suggest that mostly engagement might play a key role in the high user star rating that is contrary to previous results highlighting the impact of functionality [26,29]. Nevertheless, user star ratings should be interpreted with caution as end users lack the qualification to assess information quality. Furthermore, user star ratings lack credibility as they could originate from fictitious persons or they could refer to previous versions of the app [36].

The only relevant predictor for overall app quality was the number of functions, which is in line with previous results for apps reviews aiming at weight management, diet, physical activity, and sedentary behavior [16,22,29]. This association needs to be addressed in future studies, as it is highly likely that not only the number of functions but also their quality is crucial for overall quality. Furthermore, there might be an optimal number of functions; too many functions might be overwhelming, especially for inexperienced users. Owing to the lack of identified randomized controlled trials, no conclusions about the relationship between quality/functions and effectiveness or side effects can be drawn.

Limitations and Future Research

In this review, apps were only searched in the Google Play store and App Store and, thus, this review does not cover all the available apps promoting physical activity. However, 90% of all the apps are downloaded in these stores [45]. Owing to the broad search strategy and no focus on only the most popular apps or a cut-off concerning user ratings, this review, including 312 apps, can be seen as comprehensive. The search terms were selected after a discussion between psychologists rather than sport scientists. However, EMM is a state-certified coach (Federal Sports Academy Austria). Furthermore, EMM and HB are experts in the field of app ratings with the MARS. The stated search terms provided a more comprehensive quality analysis of apps for physical activity than previous reviews (range 13-167) [15-18]. Since new apps are being developed rapidly and the content of existing apps might have changed, the presented results can only be seen as a snapshot of the current state of the offered apps. Meanwhile, several new apps may be available and some of the included apps may be unavailable by now or may have been updated. Furthermore, apps were not tested for several hours or days. Thus, some features may not have been discovered, and some obstacles may have reminded hidden. Previous reviews evaluating apps for physical activity [15,16,29] assessed BCTs that are common to many health behavior theories [21]. Even though some of these BCTs were included in the content and functions of this review (eg, self-monitoring, feedback on performance, goal setting, or provision of information), a comparison to results concerning BCTs of previous studies is limited. The comparability between systematic app reviews in different health domains should be enhanced by using the functions included in the MARS [38,39,44]. In this systematic review, privacy and security measures were assessed on a descriptive level. Data security and privacy information were only checked based on information within the app. An in-depth analysis of privacy and security features and more elaborated strategies (eg, evaluating whether data collection and transfer are conducted according to privacy policies) are needed [28]. Future studies should extend the findings of this study by using such procedures. Lastly, it should be highlighted that the regression analyses in this study were exploratory. Thus, the results should be interpreted carefully. Confirmatory studies with adequate study designs and power are needed to identify features that are crucial elements for high app quality. Looking at the number of functions—as specified in the exploratory analyses in this study—or investigating persuasive design might be promising to begin with [46,47].

Conclusions

There is a wide range of apps offered to foster physical activity and they show overall moderate quality. High-quality apps have been presented in Multimedia Appendix 1. However, users should be aware of the broad quality range, the lack of evidence, and low ratings in privacy and security features. Thus, recommendations for the use of physical activity apps can only be given with major limitations. The contents and functions correlated positively with quality ratings. Furthermore, user ratings showed small correlations to the quality ratings and might be a limited indicator for end users. However, it seems necessary that developers use evidence-based content and scientifically developed and evaluated apps find their way into the app stores. Since the field of mobile health is rapidly growing, there is a need for continuous up-to-date evaluations of apps to provide and inform end users about data protection, privacy regulations, and evidence base. Central databases such as digital apothecaries [48-50] could help the user find high-quality apps and be protected against misinformation and abuse. However, there is also a need for novel methodological frameworks such as continuous evaluations [51] that allow for the assessment of multiple or evolving app versions.

Acknowledgments

We want to thank Julia Schmidt, Lisa Stephan, Sophie Meyer-Witting, Maximilian Schmidt, and Milena Engelke for assisting in this project. This study was self-funded (University of Ulm). The article processing charge was funded by the Baden-Wuerttemberg Ministry of Science, Research and Art and the Albert-Ludwigs-University Freiburg in the funding program

Open Access Publishing. The data sets used and analyzed in this study are available from the corresponding author on reasonable request.

Authors' Contributions

LBS, YT, HB, and EMM designed the trial and initiated this study. EMM, YT, SP, LS, JL, SC, SB, LVK, AMK, DS, SS, HB, KP, RP, and RW conducted the MARS ratings. EMM supervised all the MARS ratings. YT conducted the statistical analyses. SP wrote the first draft. All authors revised the manuscript, read, and approved the final report.

Conflicts of Interest

HB, LBS, and EMM received payments for talks and workshops in the context of e-mental health. YT, HB, LBS, and EMM co-developed and run the German Mobile Health App Database project (MHAD). The MHAD is a self-funded project at Ulm University with no commercial interests. All other authors declare that they have no competing interests.

Multimedia Appendix 1
Top 10 apps with the highest quality ratings.
[DOCX File , 20 KB - mhealth_v9i6e22587_app1.docx]

References

1. The world health report 2002: Reducing risks, promoting healthy life. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/42510/WHR_2002.pdf?sequence=1 [accessed 2021-05-06]
2. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. The Lancet 2012 Jul;380(9838):219-229. [doi: 10.1016/S0140-6736(12)61031-9]
3. Global recommendations on physical activity for health. World Health Organization. URL: https://www.who.int/dietphysicalactivity/global-PA-recs-2010.pdf [accessed 2019-08-23]
4. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. The Lancet 2012 Jul;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1]
5. Health and development through physical activity and sport. World Health Organization. URL: https://www.sportanddev.org/sites/default/files/downloads/77__health_and_development_through_physical_activity_and_sport.pdf [accessed 2021-05-06]
6. Galper DI, Trivedi MH, Barlow CE, Dunn AL, Kampert JB. Inverse association between physical inactivity and mental health in men and women. Med Sci Sports Exerc 2006 Jan;38(1):173-178. [doi: 10.1249/01.mss.0000180883.32116.28] [Medline: 16394971]
7. Physical activity guidelines advisory committee scientific report 2018. U.S. Department of Health and Human Services. URL: https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf [accessed 2019-08-23]
8. Kuhn B, Amelung V. Kapitel 4. Gesundheits-Apps und besondere Herausforderungen. In: Albrecht UV, editor. Chancen und Risiken von Gesundheits-Apps (CHARISMHA). Hannover, Germany: Medizinische Hochschule Hannover; 2016:100-114.
9. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: A meta-analysis. Behavior Modification 2015;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
10. Krebs P, Duncan DT. Health app use among US mobile phone owners: A national survey. JMIR mHealth and uHealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
11. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. Journal of Medical Internet Research 2017;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
12. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Smith S. A review of smartphone applications for promoting physical activity. Jacobs Journal of Community Medicine 2016;2(1):21. [Medline: 27034992]
13. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. Journal of Medical Internet Research 2019;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
14. Mateo GF, 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. Journal of Medical Internet Research 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
15. 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. International Journal of Behavioral Nutrition and Physical Activity 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]

16. 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;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
17. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. American Journal of Preventive Medicine 2014;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
18. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: A systematic content analysis. Journal of Medical Internet Research 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
19. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: Review and content analysis. JMIR mHealth and uHealth 2018;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
20. Brannon EE, Cushing CC. A systematic review: Is there an app for that? Translational science of pediatric behavior change for physical activity and dietary interventions. Journal of Pediatric Psychology 2015;40(4):373-384. [doi: 10.1093/jpepsy/jsu108] [Medline: 25502745]
21. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychology 2008;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
22. 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. International Journal of Behavioral Nutrition and Physical Activity 2017;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
23. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: A new tool for assessing the quality of health mobile apps. JMIR mHealth and uHealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
24. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: A systematic review. Journal of Telemedicine and Telecare 2018;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
25. Albrecht UV, editor. Chancen und Risiken von Gesundheits-Apps (CHARISMHA). Hannover, Germany: Medizinische Hochschule Hannover; 2016.
26. Armstrong S. Which app should I use? Patients and doctors are making increasing use of health apps, but there is little guidance about how well they work. British Journal of Sports Medicine 2017;51(16):1237-1239. [doi: 10.1136/bjsports-2016-g8027rep] [Medline: 28760947]
27. Grundy Q, Chiu K, Bero L. Commercialization of user data by developers of medicines-related apps: A content analysis. Journal of General Internal Medicine 2019;34(12):2833-2841 [FREE Full text] [doi: 10.1007/s11606-019-05214-0] [Medline: 31529374]
28. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. JAMA Network Open 2019;2(4):e192542 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2542] [Medline: 31002321]
29. 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. International Journal of Behavioral Nutrition and Physical Activity 2016;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
30. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One 2020;15(11):e0241480 [FREE Full text] [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
31. Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German version of the Mobile App Rating Scale (MARS-G): Development and validation study. JMIR mHealth and uHealth 2020;8(3):e14479 [FREE Full text] [doi: 10.2196/14479] [Medline: 32217504]
32. Features of the rating scale. Mobile App Rating Scale-German version. URL: https://www.youtube.com/watch?v=5vwMiCWC0Sc [accessed 2021-05-10]
33. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine 2016;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
34. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
35. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979. URL: https://www.ime.usp.br/~abe/lista/pdf4R8xPVzCnX.pdf [accessed 2021-05-14]
36. Portenhauser AA, Terhorst Y, Schultchen D, Sander LB, Denkinger MD, Stach M, et al. Mobile Apps for Older Adults: Systematic Search and Evaluation Within Online Stores. JMIR Aging 2021 Feb 19;4(1):e23313 [FREE Full text] [doi: 10.2196/23313] [Medline: 33605884]

37. Schultchen D, Terhorst Y, Holderied T, Stach M, Messner EM, Baumeister H, et al. Stay present with your phone: A systematic review and standardized rating of mindfulness apps in european app stores. International Journal of Behavioral Medicine 2020:1-9. [doi: 10.1007/s12529-020-09944-y] [Medline: 33215348]
38. Sander LB, Schorndanner J, Terhorst Y, Spanhel K, Pryss R, Baumeister H, et al. 'Help for trauma from the app stores?' A systematic review and standardised rating of apps for Post-Traumatic Stress Disorder (PTSD). Eur J Psychotraumatol 2020;11(1):1701788 [FREE Full text] [doi: 10.1080/20008198.2019.1701788] [Medline: 32002136]
39. Terhorst Y, Rathner EM, Baumeister H, Sander LB. «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]
40. Knitza J, Tascilar K, Messner EM, 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 and uHealth 2019;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
41. van Reijen M, Vriend I, van Mechelen W, Verhagen EA. Preventing recurrent ankle sprains: Is the use of an app more cost-effective than a printed booklet? Results of a RCT. Scandinavian Journal Medicine & Science in Sports 2018;28(2):641-648. [doi: 10.1111/sms.12915] [Medline: 28543566]
42. Edney SM, Olds TS, Ryan JC, Vandelanotte C, Plotnikoff RC, Curtis RG, et al. A social networking and gamified app to increase physical activity: Cluster RCT. American Journal of Preventive Medicine 2020;58(2):e51-e62. [doi: 10.1016/j.amepre.2019.09.009] [Medline: 31959326]
43. Number of apps available in leading app stores as of 4th quarter 2020. Statista. URL: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2021-05-05]
44. Terhorst Y, Messner EM, Schultchen D, Paganini S, Portenhauser A, Eder AE, et al. Systematic evaluation of content and quality of English and German pain apps in European app stores. Internet Interv 2021 Apr;24:100376 [FREE Full text] [doi: 10.1016/j.invent.2021.100376] [Medline: 33718002]
45. Mobile operating system market share worldwide 1999-2020. StatCounter. URL: https://gs.statcounter.com/os-market-share/mobile/worldwide [accessed 2020-01-09]
46. Baumeister H, Kraft R, Baumel A, Pryss R, Messner EM. Persuasive e-health design for behavior change. In: Baumeister H, Montag C, editors. Digital Phenotyping and Mobile Sensing: New developments in Psychoinformatics. Berlin, Germany: Springer; 2019:261-276.
47. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: Examining which aspects of intervention design matter most. Translational Behavioral Medicine 2018;8(5):793-798. [doi: 10.1093/tbm/ibx037] [Medline: 29471424]
48. Muñoz RF, Chavira DA, Himle JA, Koerner K, Muroff J, Reynolds J, et al. Digital apothecaries: A vision for making health care interventions accessible worldwide. Mhealth 2018;4:18 [FREE Full text] [doi: 10.21037/mhealth.2018.05.04] [Medline: 30050914]
49. Apps and digital health resources reviewed by experts. One Mind PsyberGuide. URL: https://onemindpsyberguide.org/ [accessed 2021-05-10]
50. kvappradar: Doctors and psychotherapists rate health apps. Central Institute for Statutory Health Insurance in the Federal Republic of Germany. URL: https://www.kvappradar.de/ [accessed 2021-05-10]
51. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. General Hospital Psychiatry 2013;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]

Abbreviations

BCT: behavior change technique
ICC: intraclass correlation coefficient
MARS-G: German version of the Mobile App Rating Scale

Edited by L Buis; submitted 17.07.20; peer-reviewed by C Cheng, M Alarifi; comments to author 05.10.20; revised version received 30.11.20; accepted 05.01.21; published 09.06.21.

Please cite as:
Paganini S, Terhorst Y, Sander LB, Catic S, Balci S, Küchler AM, Schultchen D, Plaumann K, Sturmbauer S, Krämer LV, Lin J, Wurst R, Pryss R, Baumeister H, Messner EM
Quality of Physical Activity Apps: Systematic Search in App Stores and Content Analysis
JMIR Mhealth Uhealth 2021;9(6):e22587
URL: https://mhealth.jmir.org/2021/6/e22587
doi:10.2196/22587
PMID:34106073

Original Paper

Mobile Apps for Drug–Drug Interaction Checks in Chinese App Stores: Systematic Review and Content Analysis

Chunying Shen[1*], MSc; Bin Jiang[2*], PhD; Qilian Yang[1*], MSc; Chengnan Wang[3], BSc; Kevin Z Lu[4], BPharm, PhD; Meng Gu[3], BSc; Jing Yuan[1*], BPharm, PhD

[1]Department of Pharmacy, Minhang Hospital, Fudan University, Shanghai, China

[2]Department of Pharmacy Administration & Clinical Pharmacy, School of Pharmaceutical Sciences, Peking University, Beijing, China

[3]Department of Clinical Pharmacy & Pharmacy Practices, School of Pharmacy, Fudan University, Shanghai, China

[4]University of South Carolina College of Pharmacy, Columbia, SC, United States

[*]these authors contributed equally

Corresponding Author:
Jing Yuan, BPharm, PhD
Department of Pharmacy
Minhang Hospital
Fudan University
826 Zhangheng Rd
Shanghai, 201203
China
Phone: 86 215 1980025
Fax: 86 213 4228568
Email: jyuan@fudan.edu.cn

Abstract

Background: As a computerized drug–drug interaction (DDI) alert system has not been widely implemented in China, health care providers are relying on mobile health (mHealth) apps as references for checking drug information, including DDIs.

Objective: The main objective of this study was to evaluate the quality and content of mHealth apps supporting DDI checking in Chinese app stores.

Methods: A systematic review was carried out in November 2020 to identify mHealth apps providing DDI checking in both Chinese iOS and Android platforms. We extracted the apps' general information (including the developer, operating system, costs, release date, size, number of downloads, and average rating), scientific or clinical basis, and accountability, based on a multidimensional framework for evaluation of apps. The quality of mHealth apps was evaluated by using the Mobile App Rating Scale (MARS). Descriptive statistics, including numbers and percentages, were calculated to describe the characteristics of the apps. For each app selected for evaluation, the section-specific MARS scores were calculated by taking the arithmetic mean, while the overall MARS score was described as the arithmetic mean of the section scores. In addition, the Cohen kappa (κ) statistic was used to evaluate the interrater agreement.

Results: A total of 7 apps met the selection criteria, and only 3 included citations. The average rating score for Android apps was 3.5, with a minimum of 1.0 and a maximum of 4.9, while the average rating score for iOS apps was 4.7, with a minimum of 4.2 and a maximum of 4.9. The mean MARS score was 3.69 out of 5 (95% CI 3.34-4.04), with the lowest score of 1.96 for Medication Guidelines and the highest score of 4.27 for MCDEX mobile. The greatest variation was observed in the information section, which ranged from 1.41 to 4.60. The functionality section showed the highest mean score of 4.05 (95% CI 3.71-4.40), whereas the engagement section resulted in the lowest average score of 3.16 (95% CI 2.81-3.51). For the information quality section, which was the focus of this analysis, the average score was 3.42, with the MCDEX mobile app having the highest score of 4.6 and the Medication Guidelines app having the lowest score of 1.9. For the overall MARS score, the Cohen interrater κ was 0.354 (95% CI 0.236-0.473), the Fleiss κ was 0.353 (95% CI, 0.234-0.472), and the Krippendorff α was 0.356 (95% CI 0.237-0.475).

Conclusions: This study systematically reviewed the mHealth apps in China with a DDI check feature. The majority of investigated apps demonstrated high quality with accurate and comprehensive information on DDIs. However, a few of the apps

that had a massive number of downloads in the Chinese market provided incorrect information. Given these apps might be used by health care providers for checking potential DDIs, this creates a substantial threat to patient safety.

(JMIR Mhealth Uhealth 2021;9(6):e26262) doi:10.2196/26262

KEYWORDS

drug interaction; MARS; app; drug safety; drugs; mHealth

Introduction

Medications are generally safe when used appropriately, but there are risks associated with medication use. Adverse drug event (ADEs), defined as injuries caused by the use of a drug [1], have emerged as a serious public health problem [2]. Every year in the United States, approximately 1 million emergency department visits and 125,000 hospital stays are related to ADEs [3]. Even though detailed data in China are largely lacking, the adverse outcomes caused by drug–drug interactions (DDIs) in China have been estimated to be more serious compared to those in other developed countries, such as the United States. An ADE can be related to a medication error, a DDI, or an adverse drug reaction. As the major contributor to ADEs [1,4], DDIs occur when one drug interferes with another [5], resulting in an altered clinical effect of one or both drugs. DDIs are associated with harmful outcomes, including hospitalizations, emergency department visits, and even death [6,7].

DDIs are avoidable, however, and preventing DDI remains a patient safety challenge in many countries including China. It has been widely reported that physicians and pharmacists, who are in the front line of detecting DDIs, cannot recognize clinically important DDIs [8]. With 20,000 drugs being approved every year, information on more than 100,000 potential DDIs is available in the major medication knowledge bases, such as Micromedex and Lexi-Interact. It is impossible for health care professionals to remember and identify all DDIs. In recent years, medication-related clinical decision support systems have been developed and implemented in some countries to detect and avoid potential DDIs. However, such a system is not common in clinical practice in China. Furthermore, medication knowledge bases often request fees and cannot be easily accessed by Chinese health care professionals, especially those practicing in rural areas.

With the advance of smartphones and mobile apps, using mobile health (mHealth) apps with a DDI checking function seems promising, especially with consideration to the convenience of searching for drug information. However, mHealth apps supporting DDI checks are not subject to the National Medical Products Administration (NMPA) regulations, posing a substantial threat to patient safety. A Canadian study recently reported the results of an assessment evaluating mHealth apps supporting DDI checks and found a lack of high-quality apps [9]. Additionally, the comprehensiveness and quality of app contents underpinning the best available evidence are rarely assessed, especially in non–English-speaking countries [10].

To our best knowledge, there is no published study evaluating the DDI-related mHealth apps available in Chinese app stores. As using incorrect drug information can have serious consequences, the aim of this study was to systematically evaluate DDI-related mHealth apps in Chinese app stores.

Methods

This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) systematic review protocol [11,12].

Search Strategy

We used the "crawling" method to interact directly with app stores' mHealth repository to avoid any personalized search results that might have been determined by a previous search query. Compared to traditional methods using the search query, creating a health-related app repository allowed us to perform a more thorough and reliable search [10,13]. The crawling method has been applied in searching for weight management [14] and nutrition-related apps [15]. First, we created an app repository by crawling all apps from the health and fitness category in the Chinese iTunes app store webpage (Apple Inc) [16] and the health category in the Chinese Android app store webpage (Google) [17] on November15, 2020. We then extracted the detailed information for the apps that were crawled from the app stores, including the app name, description, user rating, and the number of downloads [16]. Two raters, CW and GM, screened the description of apps independently to select those that supported DDI checks.

To avoid potential omissions, we also carried out an extensive keyword search in both iOS and Android app stores. The search terms included "drug interaction*," "pill interaction*," "medication interaction*," and "DDI*," following previous app review studies [9]. The inclusion criteria were the following: updated after 2018 and health care professionals as the targeted users. The exclusion criteria were the following: without any DDI information, duplicate apps, and not available in Mandarin Chinese.

CW and MG independently searched in an iPhone or Android phone and selected the apps for inclusion according to the selection criteria. In situations where there was a discrepancy, a third senior rater (CYS) reviewed the app description, and a consensus was made after a thorough discussion. All 3 raters are pharmacists who are currently practicing in the hospital.

Data Extraction

We first extracted data on the general information about the apps, including the developer, operating system (iOS, Android, or both), costs (free or paid), release date, size (in megabytes), number of downloads, and average rating in the app stores. For mHealth apps supporting a DDI check, the information quality and content accountability are critical for patient safety. Hence, we also extracted specific information related to the scientific

or clinical basis and accountability [10], according to the multidimensional framework for the evaluation of apps.

The scientific or clinical basis refers to the scientific quality of the content and was evaluated by the following metrics: accordance with existing evidence (yes or no), presence of citations (yes or no), clinician involvement (yes or no), affiliation with credible organization (yes or no), and expert assessment (yes or no) [10].

Accountability relates to the credibility of the app developer and was assessed by the presence of the following information: regulatory approval (yes or no), copyright information (yes or no), date of the last update, contact information, and disclaimer [10].

Quality Assessment

To assess the different factors related to app quality, we used the Mobile App Rating Scale (MARS), a multidimensional instrument for systematically evaluating the quality of mHealth apps [18]. The MARS is a validated tool for evaluating the quality of mHealth apps [18] and has been used in a wide range of fields, such as cardiovascular disease [19], rheumatology [20], mental health, diabetes [21], pediatrics [22], and weight management [14]. MARS is a 23-item expert-based evaluation tool, consisting of multiple dimensions to evaluate different aspects of apps, including end-user engagement, functionality, aesthetics, and information quality [18]. Each question uses a 5-point Likert type scale with a range from 1 to 5 (1 indicates inadequate and 5 indicates excellent; Multimedia Appendix 1). MARS has demonstrated high internal consistency and strong interrater reliability [18]. Procedures performed in previous research informed our method for evaluating the accuracy and comprehensiveness of DDI information [9,23-27]. We used a list of 35 DDI pairs, including 28 true-positive and 7 false-positive examples as the objective example in MARS (Multimedia Appendix 2). Based on clinical pharmacists' input, we selected these drug combinations because they have significant clinical impact and are commonly used in the existing questionnaires testing clinicians' knowledge of DDIs [9,23-27]. Before rating the apps, all raters read the MARS protocol and viewed the training video, following the MARS developer's recommendations [18]. In addition, all raters reached a consensus on the evaluation of the first app.

Statistical Analysis

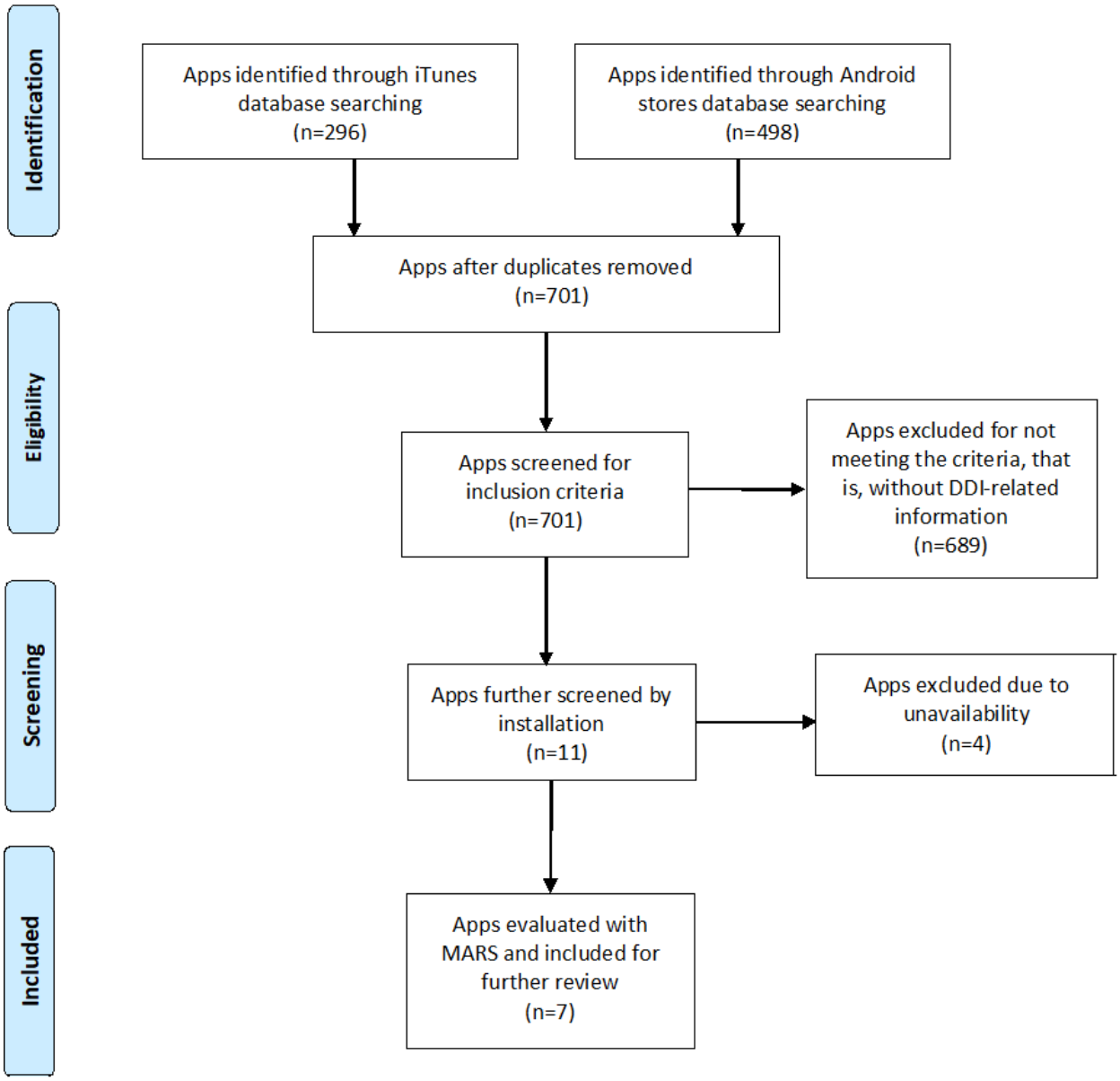
Descriptive statistics, including numbers and percentages, were calculated to describe the characteristics of the apps. For each app selected for evaluation, the section-specific MARS scores were calculated by taking the arithmetic mean, while the overall MARS score were calculated as the arithmetic mean of the section scores [18]. In addition, overall scores and section-specific MARS scores were also described by their mean, median, and IQR. The interrater agreement was examined by Cohen kappa coefficient (κ) [28] according to following scoring scheme informed by Cohen and more recent analysis [29]: $\kappa \leq 0$ indicated no agreement; 0.01-0.20 indicated poor agreement, 0.21-0.40 indicated fair agreement, 0.41-0.60 indicated moderate agreement, 0.61-0.80 indicated substantial agreement, and 0.81-1.00 indicated almost perfect agreement. Spearman correlation was calculated for the relationships among 4 sections of the MARS score, price, and number of downloads. A significance level of $P<.05$ was used in this study. All analyses were performed using SAS 9.4 (SAS Institute).

Results

Systematic Search Results

A total of 296 and 498 apps were identified in the iOS and Android App stores, respectively (Figure 1). After duplicates were removed, 701 apps remained. Of these, 689 apps were removed because they did not contain DDI information, and 4 were removed because they could not be downloaded, leaving a total of 7 apps for this evaluation.

Figure 1. Flowchart of the mobile health app selection. DDI: Drug–Drug Interaction; MARS: Mobile App Rating Scale.



Characteristics of mHealth Apps

Table 1 summarizes the general information of the apps. Out of the 7 apps included in the analysis, 6 (86%) were available free of charge. The MCDEX mobile app provides a 7-day free trial at a price of ¥89 yuan renminbi (US $13.97) per month after trial. Two apps, Medication Assistant by DXY and Medication Assistant of People's Health, offer a free version with limited access to DDI information, with a VIP subscription providing more drug information costing up to ¥30 yuan renminbi (or around US $4.71) per month. Six apps were available on both the iOS and Android platforms, and 1 app was only available on the Android platform. The average size of apps was 108.6 MB, ranging from 2.38 MB to 334.96 MB. The release dates for the Android apps were from January 29, 2012, to June 9, 2017.

Of the 6 apps with a download count available, 3 (50%) had been downloaded more than 1 million times. The average rating score for android apps was 3.5, with a minimum of 1.0 and a maximum of 4.9, while the average rating score for iOS apps was 4.7, with a minimum of 4.2 and a maximum of 4.9.

As shown in Figure 2, 3 (43%) of the 7 apps indicated sources of information (ie, citations) and provided information which was in accordance with existing evidence. Of the 7 apps, 4 (57%) indicated the involvement of clinicians in the app development, 3 (43%) had been assessed by experts in the related field, and 6 (86%) were affiliated with credible organizations. None of the apps received regulatory approval, but the MCDEX was developed under the Committee of Experts on Rational Drug Use of the National Health Commission of China, and Medication Assistant of People's Health was affiliated with the People’s Medical Publishing House Co, Ltd, which is the leading professional medical publishing company in China. The date of the latest app update ranged from March 3, 2020, to November 15, 2020. Two apps, Medication Reference and Yi Mai Tong, were developed by the same developer. The detailed information of the apps is presented in Multimedia Appendix 3.

Table 1. General information on the mobile health apps included in the review[a].

App name in English	Target market	Platform	Size (MB)	Cost per month, RMB[b] (USD)	Release date	Downloads, n[c]	Mean user rating	
							Android	iOS
MCDEX mobile	Health care professionals	iOS & Android	13.06	89 (13.97)	3/5/2015	11,193	3.4	4.2
Medication Assistant by DXY	Health care professionals	iOS & Android	169.16	30 (4.71)	11/17/2012	55,585,879	4.5	4.8
Medication Reference	Doctors, pharmacists, and other HCPs[d]	iOS & Android	334.96	free	7/10/2012	__[e]	4.9	4.5
Medication Assistant of People's Health	Doctors, pharmacists, nurses, and other HCPs	iOS & Android	86.52	12 (1.88)	6/9/2017	227,021	2.0	4.9
Medication Guidelines	Health care professionals	Android	2.38	free	—	750,760	1.0	—
DXY	Health care professionals	iOS & Android	262.67	free	1/29/2012	66,298,525	4.7	4.9
Yi Mai Tong	Health care professionals	iOS & Android	73.67	free	9/26/2013	8,900,870	4.1	4.8

[a]Apps in Chinese app stores were searched on November 15, 2020.

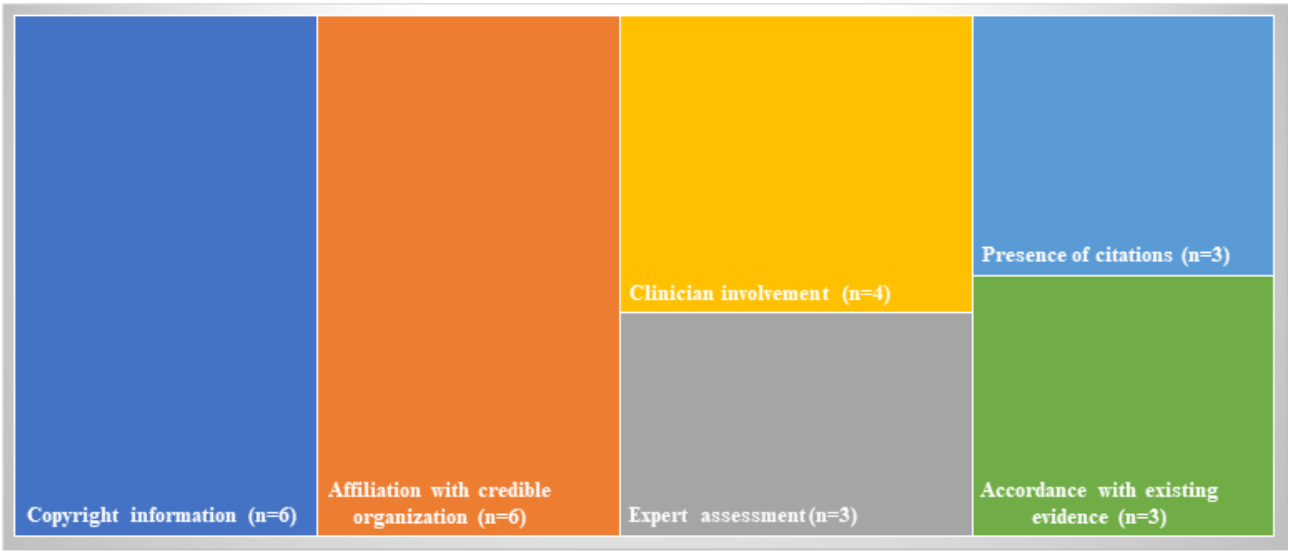
[b]RMB: yuan renminbi.

[c]Number of downloads was not available for the iOS platform.

[d]HCP: health care provider.

[e]Data not available.

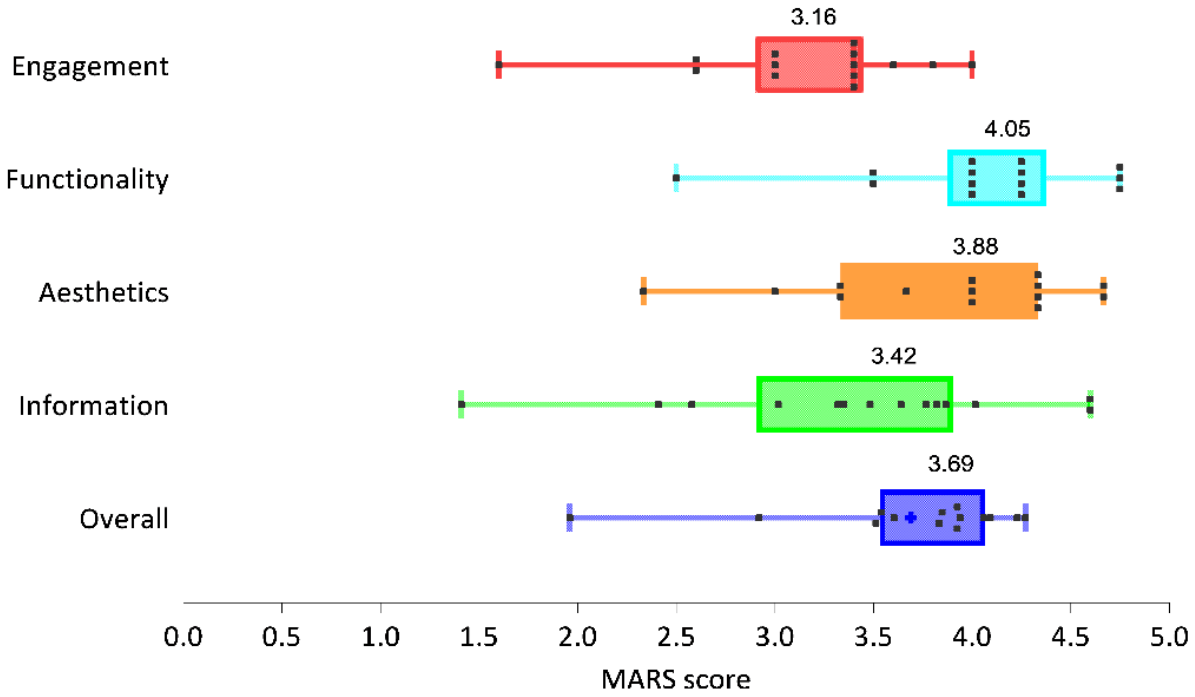
Figure 2. Information quality and accountability of the mobile health apps.



Quality Assessment

Overall, the mean MARS score was 3.69 (95% CI 3.34-4.04; Figure 3), with a lowest score of 1.96 for Medication Guidelines and a highest score of 4.27 for MCDEX mobile (Multimedia Appendix 4). A substantial variation in the MARS scores was observed among the 4 sections, with the greatest variation being observed in the information section, ranging from 1.41 to 4.60. The functionality section showed the highest mean score of 4.05 (95% CI 3.71-4.40), whereas the engagement section had the lowest average score of 3.16 (95% CI 2.81-3.51).

Figure 3. MARS scores by section. The box plot shows the mean, IQR, minimum, and maximum scores. The left and right edge of the boxes represent the first and third quartiles, the line within the boxes represents the mean, and the left and right whiskers represents the minimum and maximum scores. The scatter plot shows the distribution of MARS scores evaluated by 2 raters. MARS: Mobile App Rating Scale.



For the information quality section, which was the focus of this analysis, the average score was 3.42; the MCDEX mobile app had the highest mean score of 4.6, while the Medication Guidelines app had the lowest average score of 1.9. Of note, in the evaluation of information accuracy, the average score was 2.2 out of 5, or only 44% of the DDI pairs were described correctly. Only 1 of the 7 apps (14%), MCDEX, identified all 35 DDI pairs correctly, while 4 (57%) failed to describe half of the DDI pairs. For those DDI pairs with interactions, the average score was 2.35 out of 5, while the average score was 1.70 out of 5 for those drug pairs without DDIs, indicating these apps had a relatively higher false-positive rate. In the evaluation of the comprehensiveness of information, the average score was 2.9. Out of the 7 apps, 6 (86%) provided incomplete DDI information, while 4 apps covered less than half of the DDIs. The detailed results are presented in Multimedia Appendix 4.

For the overall MARS score, the κ coefficient was 0.354 (95% CI 0.236-0.473), the Fleiss κ was 0.353 (95% CI 0.234-0.472), and the Krippendorff α was 0.356 (95% CI 0.237-0.475). Based on the cutoff level of the κ statistic commonly cited in the literature, a κ of 0.354 was interpreted as fair agreement [29]. Multimedia Appendix 5 shows the detailed interrater reliability results.

Table 2 shows the relationship between the MARS scores and general characteristics of the 7 apps included in the analysis. There was an association between the price of mHealth apps and the MARS scores of information sections, but this did not reach statistical significance (P=.08). The number of downloads was negatively correlated with the prices of apps, even though this did not reach statistical significance. Statistically significant associations among the 4 sections of MARS scale were observed, except for the information section.

Table 2. Correlation matrix among MARS scores, price, number of downloads, and average user rating.

Characteristics	MARS[a] scores				Price	Number of downloads[b]	Mean user rating
	1	2	3	4			
MARS scores							
Engagement	—[c]						
Functionality	0.68[d]	—					
Aesthetics	0.95[e]	0.80[f]	—				
Information	0.68[g]	0.94[h]	0.82[i]	—			
Price	0.36	0.48	0.54	0.70[j]	—		
Number of downloads[b]	0.30	0.32	0.25	0.18	–0.40	—	
Average user rating	0.09	0.12	0.06	0.18	0.00	0.73[k]	—

[a]MARS: Mobile App Rating Scale.

[b]Number of downloads was not available for the iOS platform.

[c]Not applicable.

[d]*P*=.09.

[e]*P*=.001.

[f]*P*=.03.

[g]*P*=.09.

[h]*P*=.002.

[i]*P*=.02.

[j]*P*=.08.

[k]*P*=.096.

Discussion

This systematic review found there to be an acceptable quality of mHealth apps with a DDI check in Chinese app stores, with an average MARS score of 3.63. However, the quality of the information section was polarized among apps included in the review. Specifically, nearly half of the investigated apps that aimed to identify any significant interaction associated with concurrently administered drugs showed relatively poor quality in scientific information. On the other hand, the MCDEX mobile app, developed under the supervision of the China National Health Committee, demonstrated high quality in content accuracy and comprehensiveness, highlighting the importance of the scientific or clinical basis and accountability dimensions. To our best knowledge, this study was the first analysis to systematically evaluate apps for DDI checks in China and underscores the importance of regulation in the mHealth apps, which is becoming a major source of information for health care providers in China, based on our ongoing survey exploring physicians' knowledge and sources of information on DDIs.

In this in-depth analysis, the information quality of mHealth apps with a DDI check feature showed great variety. The total MARS score ranged from 1.97 to 4.23, whereas the MARS score for the information section ranged from 1.41 to 4.60, suggesting there was a certain proportion of apps with relatively low quality. These findings were consistent with another app review conducted in Canadian app stores [9]. Furthermore, 4 out of 7 investigated apps failed to identify over 50% of the tested DDI pairs. The low quality of the information section can be partially explained by a lack of evidence. The MCDEX mobile app, which was based on the knowledge bases developed in China, demonstrated relatively high-quality scientific information. A moderate score in the information section was observed in 3 apps that used both product information (package inserts) and existing treatment guidelines as the major source of information, while a poor score was observed among those apps using product information as the sole source of DDI knowledge. The reason why this is a low-quality information source is that package inserts are very likely to be outdated and cannot provide the best-available evidence. In addition, clinicians or drug experts are not commonly involved in the process of app development. The DDI check feature could be helpful in preventing ADEs but can only operate properly when the medication list is accurate and complete. Therefore, these findings call for immediate action to address the low scientific quality of mHealth apps, which is a potential threat to patient safety.

Our study also suggested that prices could be an important factor influencing the information quality of mHealth apps with DDI features. The MCDEX mobile app, which required the highest subscription fee, scored highest in the information section by providing accurate and comprehensive information in DDIs. However, 4 apps available for free provided unsatisfactory drug information. Of further note, these free apps were very popular in the market, making patient safety a serious concern.

In this review, only 7 mHealth apps were identified in Chinese app stores, fewer than those available to Canadians in English (n=26) according to a similar review conducted in 2018 [9].

This may be explained by the fact that the Canadian review targeted consumer apps with DDI check features, but the DDI check function was not included in consumer apps in China. In comparison with apps found in Canadian app stores, the Chinese apps had slightly higher MARS scores (3.63 vs 3.05) [9]. A potential explanation for this is the fact that the Chinese apps target health care professionals and thus require more rigorous clinal evidence. In terms of information quality measured by MARS, considerable variation was also observed. One-third of the Canadian apps scored lower than 1 out of 5, but around 15% (1/7) of the apps in Chinese app stores had the same score. For the functionality and aesthetic sections, similar results were reported in the Chinese and Canadian apps, but the Chinese apps scored higher in the engagement section than did the Canadian apps. If the investigated apps are mainly used as a reference for DDI information, it is not necessary to have a high score in the engagement section because these apps do not intend to engage users to effect behavioral changes [14].

This study has the following strengths. First, this is the first systematic review for DDI-related apps using the advanced crawling method for ensuring a more comprehensive search [10]. This review focused on non-English apps rather than English apps that have already been assessed [10]. Second, by using tested DDI pairs with high clinical importance, we evaluated the accuracy and comprehensiveness of the DDI checks available in the investigated apps to ensure the assessment process was as objective as possible. Finally, this review was conducted by clinical pharmacists who have extensive experience in medicine, especially DDIs.

There are also several limitations to our study. First, we searched the apps at a certain time point, and we cannot exclude the possibility that newly released apps might have been missed in the search. Second, our search might not have been sufficiently thorough. However, 2 raters performed an independent review with the consultation of a third rater, and thus the possibility that certain apps were missed should have been minimal. Third, despite the efforts to make raters familiar with the MARS scale by watching videos and reading protocols, the rating scores might have been subjective, which makes it difficult to compare across different apps. A higher score reported for a certain app may not indicate higher quality; instead, it may suggest that this app was overscored by the raters. To address this concern, we also used 35 drug pairs to assess the information quality in a more objective way.

In conclusion, this study provided a comprehensive overview of the mHealth apps with a DDI check function available in Chinese app stores. Using the multidimensional framework for the evaluation of apps, we found that the quality of mHealth apps was acceptable although a limited number of apps provided inaccurate and incomplete information about DDIs. The majority of investigated apps provided accurate and comprehensive information. A few of the apps that had large number of downloads offered a relatively low quality of drug information. As most of the apps found in Chinese app stores targeted health care professionals who may use these apps as a reference for DDI information, our findings underscore the importance of providing accurate scientific information in mHealth apps, as DDIs can have serious consequences.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mobile App Rating Scale (MARS).
[DOCX File , 23 KB - mhealth_v9i6e26262_app1.docx]

Multimedia Appendix 2
List of drug–drug interaction pairs for the Mobile App Rating Scale (MARS) number 15 and 16.
[DOCX File , 15 KB - mhealth_v9i6e26262_app2.docx]

Multimedia Appendix 3
Detailed results of information quality and accountability.
[DOCX File , 14 KB - mhealth_v9i6e26262_app3.docx]

Multimedia Appendix 4
Detailed results of Mobile App Rating Scale (MARS) number 15 and 16.
[DOCX File , 17 KB - mhealth_v9i6e26262_app4.docx]

Multimedia Appendix 5
Detailed interrater reliability results.
[DOCX File , 13 KB - mhealth_v9i6e26262_app5.docx]

References

1. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA 1995 Jul 05;274(1):29-34. [Medline: 7791255]
2. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998 Apr 15;279(15):1200-1205. [Medline: 9555760]
3. Corrigan J, Kohn LT, Donaldson MS, editors. To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press; 2000.
4. Riechelmann RP, Zimmermann C, Chin SN, Wang L, O'Carroll A, Zarinehbaf S, et al. Potential drug interactions in cancer patients receiving supportive care exclusively. J Pain Symptom Manage 2008 May;35(5):535-543 [FREE Full text] [doi: 10.1016/j.jpainsymman.2007.06.009] [Medline: 18243638]
5. Piper BJ, Alinea AA, Wroblewski JR, Graham SM, Chung DY, McCutcheon LRM, et al. A quantitative and narrative evaluation of Goodman and Gilman's pharmacological basis of therapeutics. Pharmacy (Basel) 2019 Dec 20;8(1):1 [FREE Full text] [doi: 10.3390/pharmacy8010001] [Medline: 31861770]
6. Hamilton RA, Briceland LL, Andritz MH. Frequency of hospitalization after exposure to known drug-drug interactions in a Medicaid population. Pharmacotherapy 1998;18(5):1112-1120. [Medline: 9758323]
7. Juurlink D, Mamdani M, Kopp A, Laupacis A, Redelmeier D. Drug-drug interactions among elderly patients hospitalized for drug toxicity. JAMA 2003 Apr 02;289(13):1652-1658. [doi: 10.1001/jama.289.13.1652] [Medline: 12672733]
8. Glassman PA, Simon B, Belperio P, Lanto A. Improving recognition of drug interactions: benefits and barriers to using automated drug alerts. Med Care 2002 Dec;40(12):1161-1171. [doi: 10.1097/00005650-200212000-00004] [Medline: 12458299]
9. Kim BY, Sharafoddini A, Tran N, Wen EY, Lee J. Consumer mobile apps for potential drug-drug interaction check: systematic review and content analysis using the Mobile App Rating Scale (MARS). JMIR Mhealth Uhealth 2018 Mar 28;6(3):e74 [FREE Full text] [doi: 10.2196/mhealth.8613] [Medline: 29592848]
10. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
11. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
12. 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] [Medline: 15471760]
13. Xu W, Liu Y. mHealthApps: a repository and database of mobile health apps. JMIR Mhealth Uhealth 2015 Mar 18;3(1):e28 [FREE Full text] [doi: 10.2196/mhealth.4026] [Medline: 25786060]
14. 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] [Medline: 26964880]
15. Li Y, Ding J, Wang Y, Tang C, Zhang P. Nutrition-related mobile apps in the China app store: assessment of functionality and quality. JMIR Mhealth Uhealth 2019 Jul 30;7(7):e13261 [FREE Full text] [doi: 10.2196/13261] [Medline: 31364606]
16. iTunes Store. Apple. URL: https://apps.apple.com/cn/genre/ios-%E5%8C%BB%E7%96%97/id6020 [accessed 2020-11-15]
17. Health. HUAWEI AppGallery. URL: https://appgallery.huawei.com/#/Apps [accessed 2020-11-15]
18. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
19. 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] [Medline: 27302310]
20. Knitza J, Tascilar K, Messner E, Meyer M, Vossen D, Pulla A, et al. German mobile apps in rheumatology: review and analysis using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2019 Aug 05;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
21. Wu Y, Zhou Y, Wang X, Zhang Q, Yao X, Li X, et al. A comparison of functional features in Chinese and US mobile apps for diabetes self-management: a systematic search in app stores and content analysis. JMIR Mhealth Uhealth 2019 Aug 28;7(8):e13971 [FREE Full text] [doi: 10.2196/13971] [Medline: 31464191]
22. Morse SS, Murugiah MK, Soh YC, Wong TW, Ming LC. Mobile health applications for pediatric care: review and comparison. Ther Innov Regul Sci 2018 May;52(3):383-391. [doi: 10.1177/2168479017725557] [Medline: 29714538]
23. Alrabiah Z, Alhossan A, Alghadeer SM, Wajid S, Babelghaith SD, Al-Arifi MN. Evaluation of community pharmacists' knowledge about drug-drug interaction in Central Saudi Arabia. Saudi Pharm J 2019 May;27(4):463-466 [FREE Full text] [doi: 10.1016/j.jsps.2019.01.008] [Medline: 31061613]

24. Ko Y, Malone DC, Skrepnek GH, Armstrong EP, Murphy JE, Abarca J, et al. Prescribers' knowledge of and sources of information for potential drug-drug interactions: a postal survey of US prescribers. Drug Saf 2008;31(6):525-536. [doi: 10.2165/00002018-200831060-00007] [Medline: 18484786]
25. Gilligan AM, Warholak TL, Murphy JE, Hines LE, Malone DC. Pharmacy students' retention of knowledge of drug-drug interactions. Am J Pharm Educ 2011 Aug 10;75(6):110 [FREE Full text] [doi: 10.5688/ajpe756110] [Medline: 21931448]
26. Malone DC, Abarca J, Hansten PD, Grizzle AJ, Armstrong EP, Van Bergen RC, et al. Identification of serious drug-drug interactions: results of the partnership to prevent drug-drug interactions. J Am Pharm Assoc (2003) 2004;44(2):142-151. [doi: 10.1331/154434504773062591] [Medline: 15098848]
27. Abarca J, Colon LR, Wang VS, Malone DC, Murphy JE, Armstrong EP. Evaluation of the performance of drug-drug interaction screening software in community and hospital pharmacies. J Manag Care Pharm 2006 Jun;12(5):383-389. [doi: 10.18553/jmcp.2006.12.5.383] [Medline: 16792445]
28. Kvalseth T. A coefficient of agreement for nominal scales: an asymmetric version of kappa. Educational and Psychological Measurement 2016 Sep 07;51(1):95-101. [doi: 10.1177/0013164491511008] [Medline: 5637250]
29. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]

Abbreviations

ADE: adverse drug event
DDI: drug–drug interaction
MARS: Mobile App Rating Scale
mHealth: mobile health
NMPA: National Medical Products Administration
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 05.12.20; peer-reviewed by J Ma, B Nievas Soriano; comments to author 26.12.20; revised version received 09.01.21; accepted 07.05.21; published 15.06.21.

Please cite as:
Shen C, Jiang B, Yang Q, Wang C, Lu KZ, Gu M, Yuan J
Mobile Apps for Drug–Drug Interaction Checks in Chinese App Stores: Systematic Review and Content Analysis
JMIR Mhealth Uhealth 2021;9(6):e26262
URL: https://mhealth.jmir.org/2021/6/e26262
doi:10.2196/26262
PMID:33962910

Review

mHealth Interventions for Treatment Adherence and Outcomes of Care for Cardiometabolic Disease Among Adults Living With HIV: Systematic Review

Oluwakemi Ololade Odukoya[1], MPH, MSc, MD; Chidumga Ohazurike[2], MPH, MD; Maxwell Akanbi[3], MSCI, MD, PhD; Linda C O'Dwyer[4], MA, MSLIS; Brenda Isikekpei[2], MD; Ewemade Kuteyi[2], MD; Idaomeh O Ameh[5], MD; Olanlesi Osadiaye[2], MPH, MD; Khadijat Adebayo[6], MD; Adewunmi Usinoma[2], MPH, MD; Ajoke Adewole[2], MPH, MD; Nkiruka Odunukwe[7], MD; Kola Okuyemi[8], MD; Andre Pascal Kengne[9], MD

[1]Department of Community Health and Primary Care, College of Medicine, University of Lagos, Lagos, Nigeria

[2]Department of Community Health, Lagos University Teaching Hospital, Lagos, Nigeria

[3]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[4]Galter Health Sciences Library and Learning Center, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[5]Division of Nephrology, Zenith Medical and Kidney Center, Abuja, Nigeria

[6]Department of Clinical Medicine, All Saints University School of Medicine, Roseau, Dominica

[7]Non-Communicable Disease Research Group, Nigeria Institute of Medical Research, Lagos, Nigeria

[8]Department of Family and Preventive Medicine, University of Utah School Of Medicine, Salt Lake City, UT, United States

[9]Non-Communicable Disease Research Unit, Medical Research Council, Cape Town, South Africa

Corresponding Author:
Oluwakemi Ololade Odukoya, MPH, MSc, MD
Department of Community Health and Primary Care
College of Medicine
University of Lagos
#16 Ishaga Road, Idi-Araba
Lagos, PMB 12003
Nigeria
Phone: 234 8023200770
Email: drolukemiodukoya@yahoo.com

Abstract

Background: The success of antiretroviral therapy has led to an increase in life expectancy and an associated rise in the risk of cardiometabolic diseases (CMDs) among people living with HIV.

Objective: Our aim was to conduct a systematic review to synthesize the existing literature on the patterns of use and effects of mobile health (mHealth) interventions for improving treatment adherence and outcomes of care for CMD among people living with HIV.

Methods: A systematic search of multiple databases, including PubMed-MEDLINE, Embase, CINAHL, Scopus, Web of Science, African Journals online, ClinicalTrials.gov, and the World Health Organization Global Index Medicus of peer-reviewed articles, was conducted with no date or language restrictions. Unpublished reports on mHealth interventions for treatment adherence and outcomes of care for CMD among adults living with HIV were also included in this review. Studies were included if they had at least 1 component that used an mHealth intervention to address treatment adherence or 1 or more of the stated outcomes of care for CMD among people living with HIV.

Results: Our search strategy yielded 1148 unique records. In total, 10 articles met the inclusion criteria and were included in this review. Of the 10 studies, only 4 had published results. The categories of mHealth interventions ranged from short messaging, telephone calls, and wearable devices to smartphone and desktop web-based mobile apps. Across the different categories of interventions, there were no clear patterns in terms of consistency in the use of a particular intervention, as most studies (9/10, 90%) assessed a combination of mHealth interventions. Short messaging and telephone calls were however the most common interventions. Half of the studies (5/10, 50%) reported on outcomes that were indirectly linked to CMD, and none of them provided

reliable evidence for evaluating the effectiveness of mHealth interventions for treatment adherence and outcomes of care for CMD among people living with HIV.

Conclusions: Due to the limited number of studies and the heterogeneity of interventions and outcome measures in the studies, no definitive conclusions could be drawn on the patterns of use and effects of mHealth interventions for treatment adherence and outcomes of care for CMD among people living with HIV. We therefore recommend that future trials should focus on standardized outcomes for CMD. We also suggest that future studies should consider having a longer follow-up period in order to determine the long-term effects of mHealth interventions on CMD outcomes for people living with HIV.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42018086940; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018086940

(JMIR Mhealth Uhealth 2021;9(6):e20330) doi:10.2196/20330

KEYWORDS

mHealth; HIV; cardiometabolic disease; text messaging; mobile; systematic review; telephone calls; wearable devices; smartphones; desktop; web-based; mobile apps

Introduction

Cardiometabolic diseases (CMDs) represent a huge threat to the global progress that has been achieved in reducing mortality and morbidity among people living with HIV [1,2]. With the successes of antiretroviral therapies (ARTs), there has been increased life expectancy and a reduction in the burden of opportunistic infections among people living with HIV [3-5]. However, both HIV and ART are independently associated with an increased risk of CMD [6-8]. The incidence of CMD among people living with HIV ranges from 1.19 per 1000 person-years to 11.3 per 1000 person-years [9,10]. People living with HIV show consistent patterns of increased risk for diabetes [11-13], stroke [14], sudden cardiac death [15], heart failure [16], coronary heart disease, and myocardial Infarction [17-19]. Furthermore, established CMD risk factors like tobacco smoking and alcohol use are prevalent among people living with HIV [20,21].

The use of mobile technology, especially the use of mobile phones, has increased tremendously worldwide. The majority of the over 7 billion mobile phone users reside in low- and middle-income countries (LMICs), where the burden of HIV and AIDS is the heaviest [22]. The improved access to mobile technology presents an immense opportunity for promoting the health of people living with HIV, and mobile technology is increasingly being used to provide support for people living with HIV and promote adherence to ART [23-25].

Research on CMD among people living with HIV has increased significantly over the last 2 decades, and significant proportions of these studies have assessed the use of mobile health (mHealth) interventions to promote HIV care [7-12]. Systematic reviews on the topic have shown that most research in this field focuses on mobile phone interventions that aim to promote ART adherence and other forms of direct HIV and AIDS outcomes [25-30]. In a paper that systematically reviewed mobile phone SMS text messaging interventions for HIV and other chronic diseases, many of the studies with chronic disease outcomes did not include people living with HIV [30]. Thus, the evidence for analyzing the patterns of use and effects of mHealth interventions for CMD outcomes among people living with HIV remains unclear. This review aims to synthesize existing literature on the patterns of use and effects of mHealth interventions for treatment adherence and outcomes of care for CMD among adults living with HIV.

Methods

Study Design

For this review, we followed the guidelines from the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) Statement [31]. The study protocol was also registered with the International Prospective Register of Systematic Reviews (trial ID number: CRD42018086940). We relied on data obtained from other studies that conducted primary research; therefore, obtaining institutional ethical approval was not required. We made use of deidentified data that were stored in a password-protected database. We ensured that the data presented in this systematic review did not violate the privacy of the patients.

Eligibility Criteria

Inclusion Criteria

To ensure that we captured all articles that described patterns of existing mHealth interventions for CMD outcomes, we did not limit our search by study type or design. We included studies that described any mHealth interventions if the following criteria were met: (1) the study was conducted among adults aged ≥18 years living with HIV; (2) the study involved the use of mobile phones or had any mHealth components embedded in its design; and (3) the study was designed to influence adherence to treatment or outcomes of care for 1 or more CMD (ie, hypertension, dyslipidemia, obesity, stroke, coronary heart disease, diabetes mellitus, and metabolic syndrome). Articles published in any language with English abstracts were eligible for inclusion.

Exclusion Criteria

A study was excluded if (1) it was an opinion piece, (2) it was a publication that lacked primary data, and (3) it had no explicit method description. In the case of duplicate publications of the same material in more than 1 journal or conference proceeding, the most complete and recent version was used.

Definitions

The diagnosis of HIV was based on positive reports from screening tests or participants' self-reported doctor diagnoses. The CMDs reviewed included hypertension, diabetes mellitus, dyslipidemia, obesity, stroke, coronary heart disease, and metabolic syndrome. This study was a systematic review for which no primary data were collected. The diagnosis of HIV and cardiometabolic outcomes were based on reports from the studies included in this review.

The outcome measures included differences in adherence to treatment or outcomes relating to any of the listed CMDs; differences in mortality due to the listed CMDs; differences in blood pressure, glycemic control, and blood lipid levels; and reductions in CMD risk or BMI waist circumference and waist-hip ratios.

With regard to mHealth, for the purpose of this review, we used the Global Observatory for eHealth definition, which defines mHealth as a medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, PDAs, and other wireless devices. mHealth involves the use and capitalization of a mobile phone's core utility in terms of voice messaging services and SMSs as well as more complex functionalities and applications, including general packet radio services, third and fourth generation mobile telecommunications (3G and 4G systems), GPSs, and Bluetooth technology [32].

Search Strategy for the Identification of Relevant Studies

Electronic searches of the following databases were conducted from inception to September 2019: PubMed-MEDLINE, Embase, CINAHL, Scopus, Web of Science, Cochrane Central Register of Controlled Trials, Global Health (Elton B. Stephens Company), the Institute of Electrical and Electronics Engineers, African Journals online, the Association for Computing Machinery, World Health Organization (WHO) reports, ClinicalTrials.gov, The Pan African Clinical Trials Registry and mHealth alliance, and the WHO Global Index Medicus. The search terms used included HIV-related terms (eg, *HIV infections*, *HIV*, *HIV/AIDS*, *HIV-positive*, and *HIV infected*), mobile device–related terms (eg, *mobile health*, *mHealth*, *mobile phone*, and *short message*) and CMD-related terms (eg, *Hypertension*, *Diabetes*, *stroke*, *metabolic syndrome*, and *cardiometabolic*). No language, publication type, or date limits were applied to the initial searches. The first author (OOO) and the librarian (LO) collaboratively developed the search strategy. The full list of search strategies is available in Multimedia Appendix 1.

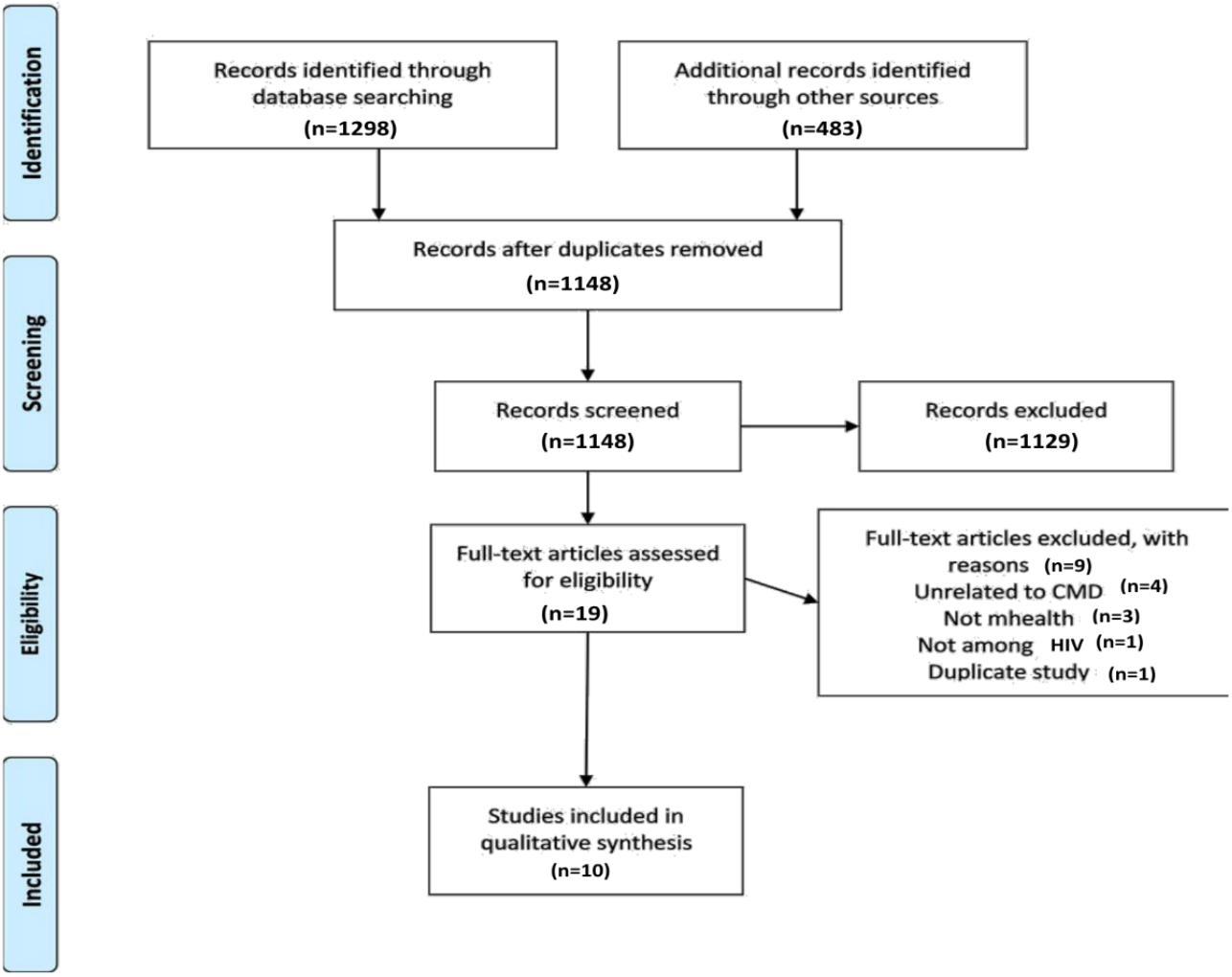
Reference Lists and Grey Literature

We searched for additional relevant articles in the reference lists of the retrieved key articles and reviews. We contacted authors of included studies to acquire other data related to our outcomes of interest that may have been unpublished, informally published, or undergoing ongoing analysis. Principal investigators of registered clinical trials whose published results were not found within our search were contacted by email to share their results.

Data Collection and Processing

Search results were saved into EndNote (Clarivate Analytics) files by the librarian (LO). All EndNote files were deduplicated, collated, and transferred into Rayyan (Rayyan Systems Incorporated) [33] for subsequent processing. A pilot screen of 100 articles was performed by each reviewer to ensure the consistent interpretation of the inclusion and exclusion criteria. Two sets of reviewers (set 1: BI and KA; set 2: OO and OU) conducted an initial independent screening and eligibility assessment of articles' titles and abstracts by using the predefined inclusion and exclusion criteria. A third reviewer (OOO) resolved disagreements. Full-text copies of the selected articles were obtained for further review and assessed using the same process as the title and abstract screens. The flowchart for the study selection process is shown in Figure 1. Two independent reviewers (BI and OO) used a pretested data extraction form that was adapted from the Cochrane data extraction template for intervention reviews of randomized controlled trials (RCTs) and non-RCTs [34] (Multimedia Appendix 2) to extract the data from the full texts [35-38]. The information extracted included the study title, author, year, country, study design, sample size, study population and setting, intervention type and delivery, components of the intervention, concurrent non-mHealth interventions or medications (if any), duration of intervention, cardiometabolic outcome, and secondary outcome measures. The results were synthesized and presented as a narrative synthesis of the details regarding the study type, intervention characteristics, study outcomes, and location. Due to the limited number of studies and the wide variability in the outcome measures, we were unable to perform a meta-analysis.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. CMD: cardiometabolic disease.



Methodological Quality of Included Studies

The quality of included studies was assessed by using the Cochrane risk of bias assessment tool [39]. Specifically, we assessed the quality of each included study by using the criteria from the seven domains of the tool (ie, random sequence generation, the blinding of study participants and key personnel, the blinding of the outcome assessment, selective outcome reporting, allocation concealment, incomplete outcome data, and the presence of bias from other sources) [39]. For each of these domains, each study was identified as having a low, high, or unclear risk. In total, 2 studies did not report on the blinding of the outcome assessment [35,36], while 3 studies had incomplete outcome data [35,36,38]. Two investigators (BI and OO) independently assessed the quality of included studies, and a third reviewer (OOO) resolved discrepancies.

Results

Included Studies

Our search yielded 1148 unique records, of which 1129 were excluded for the reasons specified in Figure 1. Most of the excluded records (1129/1148, 98.34) were not related to mHealth, not related to our CMD outcomes of interest, or not conducted among people living with HIV. Of the 19 articles that were fully reviewed, 4 were unrelated to the CMD outcomes of interest, 3 were not about mHealth interventions, 1 was not conducted among people living with HIV, and 1 was a duplicate study. In total, 10 articles met the inclusion criteria.

The 10 included studies and their key study characteristics are summarized in Table 1. Of these included studies, 7 were RCTs [35,36,40-43]. In 90% (9/10) of studies, study participants only included people living with HIV. The remaining 10% of studies included a combination of both people living with HIV and people living without HIV. Most studies (4/10, 40%) had an average study period of about 12 months [36,37,40,43]. It is noteworthy that 6 of the 10 included studies were registered clinical trials that had unpublished results at the time of conducting this review [40-45]. The principal investigators of the included registered clinicals trials were contacted in order to ascertain the progress of the trials; however, no response was received at the time of concluding this review. The number of studies with unpublished results and the variation in outcomes across studies made it difficult to conduct a meta-analysis.

Table 1. Summary of the key characteristics of included studies (N=10).

Study (author, year, country)	Study design and methods	Inclusion criteria	Interventions	Outcomes	Reported results
Morillo-Verdugo et al [36], 2018, Spain	Randomized controlled trial	Aged >35 years; on antiretroviral therapy with at least 1 drug prescribed for the treatment of hypertension, dyslipidemia, angina pectoris, cardiovascular prophylaxis, or type 2 diabetes mellitus; and at a moderate or high risk of cardiovascular disease	Periodic text messages on mobile phones	Cardiovascular risk index, smoking reduction, blood pressure control, and medication adherence	20.7% of patients in the intervention group vs 12.5% of patients in the control group reduced their Framingham risk score from high/very high to moderate/low (*P*=.02), and the number of patients with controlled blood pressure increased by 32.1% (*P*=.01). 37.9% of patients overall stopped smoking (*P*=.001).
Anglada-Martinez et al [35], 2016, Spain	Single-arm, prospective pre- and postintervention study	Patients on treatment for heart failure, hypertension, or dyslipidemia for >1 month and those aged >18 years	Medplan smartphone app and weekly motivational messages	Medication adherence, cholesterol, triglycerides, and blood pressure control	The proportion of missed doses decreased significantly for patients using the Medplan app (*P*<.05). There was no difference in the health outcomes of patients.
Roos et al [37], 2014, South Africa	Randomized controlled trial	On antiretroviral therapy for >6 months, aged 20-65 years, ambulatory without assistive device, and had an elevated risk of ischemic heart disease	Pedometer, activity diary that included education materials and documents for self-monitoring, and 1 monthly cell phone SMS message for motivation	The pedometer step count of both groups improved significantly.	The pedometer step counts of the control and intervention groups improved significantly (*P*=.03 for both groups) at 6 months, but this improvement was not significant at 12 months (*P*=.33 and *P*=.21, respectively). Significant between-group effects were observed in 6-minute walk test distances (*P*=.01), waist-to-hip ratios (*P*<.01), glucose levels (*P*<.01), and high-density lipoprotein levels (*P*<.01) over the 12-month period.
Zuniga et al [38], 2019, United States of America	1-group pre- and posttest design	Aged >18 years and had HIV and type 2 diabetes mellitus	6-hour educational instruction implemented as 2 3-hour meetings followed byweekly telephone support calls for 6 weeks	Diabetes self-management skills and knowledge about HIV or diabetes	There was a 34% increase in diabetes self-management skills from pretest to posttest, but there were no changes in knowledge about HIV or diabetes.
Grinspoon [40], 2006, United States of America	Randomized case control study	Aged 18 to 65 years and had 3 of the following 5 characteristics: (1) waist circumferences of >102 cm (40 in) for men and >88 cm (35 in) in women; (2) triglyceride levels of ≥150 mg/dL or current antilipolytic drug treatment; (3) high-density lipoprotein levels of <40 mg/dL for men and <50 mg/dL for women; (4) blood pressure of ≥130/85 mmHg or current antihypertensive drug treatment; and (5) fasting glucose level of ≥110 mg/dL	1-time counseling session with nutrition staff at the baseline visit and monthly unscripted phone calls	Waist-hip ratios and cardiovascular indices (total cholesterol; low-density lipoprotein, high-density lipoprotein, and triglyceride cholesterol levels; blood pressure; cardiac enzymes; C-reactive protein; tissue plasminogen activator; plasminogen activator inhibitor; insulin; and glucose metabolism)	The results of the study have yet to be published.

Study (author, year, country)	Study design and methods	Inclusion criteria	Interventions	Outcomes	Reported results
Jaggers et al [41], 2013, United States of America	Randomized controlled trial	Aged >18 years, had a sedentary lifestyle, had a viral load of >75 copies/mL, was capable of performing required exercise regimen, and had daily access to a telephone for approximately 10 months	Home-based physical activity intervention: The intervention included a 60-min, individual, face-to-face session; telephone counselling calls; and educational workbooks and pedometers for the self-monitoring of physical activity.	The effect of the intervention in terms of decreasing modifiable risk factors and increasing physical activity among people living with HIV and the effect of the intervention in terms of decreasing modifiable risk factors, such as fat distribution, blood lipids, and cardiorespiratory fitness outcomes, were assessed.	The findings of the study have yet to be published.
Brooke [42], 2017, United States of America	Randomized controlled trial	People living with HIV	Personalized, automated, interactive mobile phone text message intervention	Physical activity and dietary assessments; polyunsaturated fatty acids, carotenoids, and other biomarkers in plasma; and total cholesterol, triglyceride, and high- and low-density cholesterol	The study is still ongoing.
Dodson et al [43], 2016, Australia	Cluster randomized controlled trial	Aged >30 years, was receiving care from a participating doctor, was not diagnosed with cardiovascular disease, and had not participated previously in an HIV-specific self- management or coaching program	Health map website for (1) routine clinic visits involving the sharing of health records with a doctor; (2) access to own health record and information from home; (3) access to telephone and web-based self-management support; and (4) access to a peer-moderated, web-based group chat program.	10-year risk of nonfatal acute myocardial infarction or coronary heart disease death, as estimated by a Framingham Heart Study risk equation and the Positive and Active Engagement in Life Scale from the Health Education Impact Questionnaire	The findings of the study have yet to be published.
Oduor et al [44], 2018, Kenya	Contextual user interviews	Patients living with HIV and hypertension	Integrated desktop and mobile app	Improved efficacy, safety, and personalization of medication prescription	Descriptive study
Kengne [45], 2019, South Africa	Randomized controlled trial	Adult South Africans with comorbid HIV and hypertension	Automated text messaging	Mean difference in systolic and diastolic blood pressure at baseline and follow-up, uptake and adherence to blood pressure medications, mean change in lipid variables, and mean change in adiposity variables	The results of this study have yet to be published.

All included studies had an mHealth component, although some studies used a multimethod mHealth intervention approach, wherein support telephone calls were combined with educational instructions [38,41] or the use of mobile phone apps were combined with user interviews [44]. However, short messaging [36,37,42,45] and telephone calls [38,40,41,43] were the most common mHealth interventions. Furthermore, 2 studies used mobile apps as their mHealth intervention component [35,44].

Although studies had varied outcomes, most had treatment adherence (1/10, 10%) [44], cardiometabolic outcomes (5/10, 50%) [37,40-43], or both (3/10, 30%) [35,36,45] as their primary or secondary study outcomes. The other studies had a combination of other outcomes that were not limited to treatment adherence or cardiometabolic outcomes, such as physical activity and the self-management of diabetes [38,42].

With regard to the function of mHealth interventions, 4 studies used mHealth for medication adherence purposes [35,36,44,45], 3 studies used mHealth to improve physical activity and thereby reduce cardiovascular risk [37,41,42], 2 used mHealth for health promotion and health coaching purposes to reduce cardiovascular risk [37,38], and 1 used mHealth for the self-management of diabetes mellitus [38].

The majority of the studies (9/10, 90%) were conducted in hospital or clinic settings [35-38,40-43,45], whereas 10% (1/10) of the studies were conducted in a community setting in a rural area [44]. Furthermore, most of the studies were conducted in the high-income countries of Spain (2/10, 20%), the United

States of America (4/10, 40%), and Australia (1/10, 10%). Only 3 studies were conducted in the LMICs of South Africa and Kenya in Sub-Saharan Africa, where the burden of HIV is quite significant.

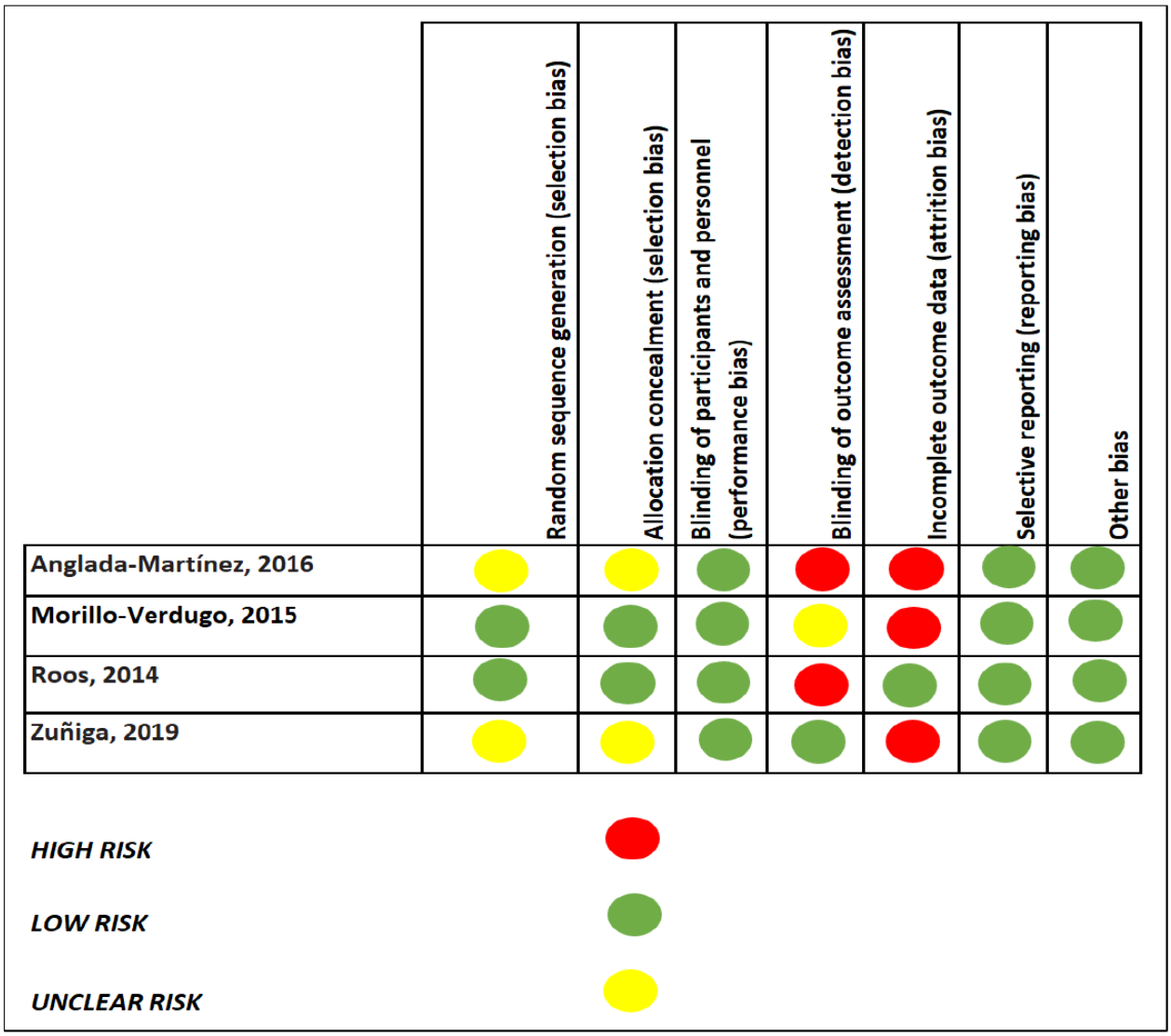
Effectiveness of mHealth Interventions

With regard to the effects of mHealth interventions on treatment adherence and cardiometabolic outcomes in the 4 studies with published results, 2 studies reported a decrease in cardiovascular risk [36,37] and significant group effects for cardiometabolic outcomes, such as those on glucose level and high-density lipoproteins, between intervention and control groups [37]. In total, 1 study reported a 34% increase in diabetes self-management skills among participants [38], and 1 study reported no differences in effects on treatment adherence among study participants [35].

Assessing the Risk Of Bias

Figure 2 shows the assessment of the risk of bias in the included studies.

Figure 2. Cochrane risk of bias assessment [35-38].



Selection Bias (Random Sequence Generation and Allocation Concealment)

Among the 10 included studies, the assessment of bias was conducted in only the 4 studies that were completed (Anglada-Martinez et al [35], Morillo-Verduogo et al [36], Roos et al [37], and Zuniga et al [38]). In the study by Morillo-Verduogo et al [36], the participants were randomized into control and intervention groups with a specific software that was used to generate a sequence of random numbers. Roos et al [37] used a simple randomization formula in Microsoft Excel 2010 for randomization. The method for selecting participants was unclear in the study by Anglada-Martinez et al [35], as the patients were followed during the preintervention phase in accordance with their usual schedule. Zuniga et al [38] selected 25 participants by using a convenient sampling method. This was because the Zuniga et al [38] study had to be completed in 12 months due to the guidelines of their funding mechanism as well as the constraints from budget limitations. During the allocation of participants, Roos et al [37] used an academic who was not directly involved in the study to carry out allocation concealment via sequentially numbered envelopes.

In the studies by Zuniga et al [38] and Anglada-Martinez et al [35], allocation concealment was not carried out for participants. This was because the Zuniga et al [38] study was a single-group study with a pre- and posttest study design, and in the Anglada-Martinez et al [35] study, the sequence generation was not described.

Performance and Detection Bias (Blinding of Participants, Personnel, and the Outcome Assessment)

Only 1 out of the 4 completed studies carried out blinding, and this was the blinding of the study personnel. In that study the research assistant performed all of the assessments and was blinded to the group allocation. The assessment forms were also coded to ensure anonymity, and the first author conducted the intervention [37].

The studies by Anglada-Martinez et al [35] and Morillo-Verduogo et al [38] did not record the blinding of participants and personnel; however, this was unlikely to influence the outcomes of those studies. The blinding of participants and personnel was not conducted in the Zuniga et al [38] study, as the graduate research assistant already knew the participants because they recruited, enrolled, and collected the baseline and follow-up data. This may likely have had an influence on the outcome of the study.

With regard to detection bias, the study by Anglada-Martinez et al [35] noted that the participants were not blinded to the measurement of the outcomes. This may have resulted in a Hawthorne effect, whereby the participants may have modified their behaviors because they were aware that their adherence to medication was being observed, thereby influencing the outcome of the study. Although the assessment forms were noted to be coded in the study by Roos et al [37], the blinding of the participants in the intervention group was not performed. These participants had 5 contact sessions in which they were educated on ischemic heart disease risk factors and the benefits of and methods for increasing physical activity levels. The participants in the control group experienced an increase in their pedometer step counts, which may have been linked to the fact that they were aware of their behavior (the Hawthorne effect). These factors were likely to influence the outcome of the Roos et al [37] study. Additionally, during the completion of the study, the intervention participants reported that wearing the pedometer motivated them to increase their activity levels [38]. The studies by Morillo-Verduogo et al [36] and Zuniga et al [38] failed to describe whether any form of blinding for the outcome assessor was performed.

Attrition and Reporting Bias (Incomplete Outcome Data and Selective Reporting)

In the study by Anglada-Martinez et al [35], there was a loss to follow-up of 20 participants. They were lost as a result of changing their addresses (n=3), transitioning to the use of a non-Android or non-iOS device (n=2), having incompatible Android and iOS operating systems (n=6), dying (n=1), not attending follow-ups (n=7), and withdrawing from the study (n=1). It was also noted that the adherence rates recorded with the use of the app did not correlate with the adherence rates that were recorded by analyzing the proportion of days in which participants took their medication. Some participants experienced problems such as their reminders not working properly (50%), and some believed that the use of the app resulted in extra work (fourth month: 60.2%; this decreased to 56.6% in the sixth month). Further, pending Medplan alerts, which appear during the day, disappear at midnight; hence, patients who took their medication before midnight were unable to confirm that they had done so after midnight (57.1%) [35]. The prespecified outcomes of interest were however reported in accordance with the study protocols [35].

In the study by Morillo-Verdugo et al [36], there was loss to follow-up of 5 participants, and 1 died from causes that were beyond the scope of the study. It was also noted that because of the low number of patients with diabetes in the study, glycosylated hemoglobin was not included in the analysis. The prespecified outcomes of interest, such as reductions in cardiovascular risk indices based on Framingham scores, adherence to ART, and lifestyle modifications, were reported in accordance with the study protocol [36].

In the study by Roos et al [37], there were losses to follow-up in the control group at 6 months (n=8) and 12 months (n=10) and in the intervention group at 6 months (n=3) and 12 months (n=6). In the control group, 3 participants did not return for their second baseline assessment, and their baseline blood results were managed by imputing the mean value of the cohort's results. Additionally, some data were noted to be missing completely at random, and these were managed by imputing the last observation that was carried forward [37]. The prespecified outcomes of interest, such as perceived stress, physiological measures, physical activity, physical function capacity, biochemical measures, and Framingham risk scores, were reported in accordance with the study protocol [37].

In the study by Zuniga et al [38], 15 out of 25 patients participated in the baseline fasting blood tests, but only 1 completed the blood drawing for the collection of follow-up data. Therefore, the effects of the intervention on HIV or diabetes control could not be recorded. The reasons for losses to follow-up were not described; however, the reporting of prespecified outcomes of interest, such as the knowledge of diabetes and HIV and diabetes self-management activities, were reported in accordance with the study protocol [38].

Discussion

Principal Findings

This paper presents a narrative synthesis of mHealth interventions for treatment adherence and outcomes of care for CMD among adults with HIV. A total of 10 studies met the inclusion criteria and were included in the review. The majority of studies included in this review were conducted in high-income countries (7/10, 70%) [35,36,38,40-43], and only a handful of studies were from LMICs (3/10, 30%) [37,44,45].

Based on our review, the categories of the interventions that were used in the studies ranged from short messaging and telephone calls to wearable devices and smartphone and desktop web-based mobile apps. However, the two most common interventions that were provided to people living with HIV were

short messaging and telephone calls. This finding is similar to those from another systematic review on the impact of mHealth chronic disease management on treatment adherence and patient outcomes, which found that 40% of the studies included in their review had used short messaging to track medication adherence in patients with chronic diseases [46]. We also found that across the different categories of interventions, there were no clear patterns in terms of consistency in the use of a particular intervention, as most studies (9/10, 90%) [35,36,38,40-45] assessed a combination of mHealth interventions.

Overall, authors reported that the use of mHealth interventions for treatment adherence and outcomes of care for CMD among adults living with HIV was effective. However, studies varied widely in terms of research questions, target groups, study outcomes, and settings [36,37,41,43,44]. The risk of bias varied from study to study. Of the 4 studies that assessed the risk of bias, 2 controlled for selection bias by randomization [35,37], and only 1 study performed the blinding of both participants and research personnel for the control of performance bias [37]. All 4 studies reported their prespecified outcomes of interest in accordance with their study protocols.

Most existing studies on mHealth interventions for people living with HIV have addressed ART adherence outcomes, and only a few have assessed mHealth interventions for CMD outcomes. In theory, the use of mHealth interventions to monitor treatment adherence and outcomes of care for both ART and CMD should make the process of care more efficient. However, research in this area is still very limited. This highlights the need to generate evidence to promote the use of integrated models of care for outcomes such as ART adherence and CMD outcomes. In addition, our findings from this review revealed that most studies did not report outcomes such as a reduction in the incidence of CMD, which should be the ultimate goal, given the increasing life expectancy of people living with HIV resulting from ART. However, this could be explained by the short follow-up periods that were used in these studies. We highlight the existing mHealth interventions that specifically target CMD outcomes among people living with HIV and draw attention to the gaps and opportunities in mHealth interventions for comorbid CMDs among people living with HIV. Furthermore, our review shows the paucity of well-designed RCTs in this research area. We also call attention to the disparities in the conduct of research on this topic. Globally, the WHO African region has been and remains to be the most severely affected by the HIV epidemic, as it accounts for more than two-thirds of the people living with HIV worldwide and nearly 3.7% adults (about 1 in every 25) living with HIV [47]. However, only 1 out of the 4 completed studies and 3 out of the 10 reviewed studies were conducted in an LMIC setting.

Although we purposefully used broad inclusion criteria to capture all studies evaluating any type of mHealth intervention for CMD among people living with HIV, we were limited by the low number of studies that met our inclusion criteria or reported on our predefined key outcomes. Furthermore, many of the studies that fit our inclusion criteria were old. Therefore, they may not have reflected the current state of the effectiveness of mHealth interventions. We also could not conduct a meta-analysis due to the heterogeneity of the included studies in terms of their methods and reported outcomes. However, it is important to point out that a number of clinical trials are underway, and their results can be incorporated in a follow-up review within the next few years. We recommend that future trials should focus on standardized outcomes for CMD to enable the conduction of a meta-analysis. We also suggest that future studies should consider using an integrated approach and a longer follow-up period in order to determine the long-term effects of mHealth interventions on outcomes of care.

Conclusion

Studies using mHealth interventions that specifically target CMD outcomes for people living with HIV are limited, particularly in Sub-Saharan Africa, where the burden of HIV is the greatest. In this review, although several of the mHealth interventions were found to be effective, there appears to be no clear pattern in the use of mHealth interventions for CMD outcomes. Short messaging was the most used intervention. More studies that assess the use and effectiveness of mHealth interventions other than short messaging, such as mobile apps and wearable health devices, are needed in this study area.

Acknowledgments

APK is supported by a European and Developing Countries Clinical Trials Partnership senior fellowship (grant TMA2017GSF-1962; Cardiometabolic Diseases Risk Evaluation and Reduction in African People Living with HIV Infection) through the South African Medical Research Council. OOO is supported by the Fogarty International Center of the National Institutes of Health under award number K43TW010704. The content of this review is solely the responsibility of the authors and does not necessarily represent the official views of the South African Medical Research Council or the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings.
[PNG File , 887 KB - mhealth_v9i6e20330_app1.png]

Multimedia Appendix 2

Data extraction template [<xref ref-type="bibr" rid="22ref35">35</xref>-<xref ref-type="bibr" rid="22ref38">38</xref>]. [PNG File , 1146 KB - mhealth_v9i6e20330_app2.png]

References

1. Trickey A, May MT, Vehreschild J, Obel N, Gill MJ, Crane H, Antiretroviral Therapy Cohort Collaboration (ART-CC). Cause-specific mortality in HIV-positive patients who survived ten years after starting antiretroviral therapy. PLoS One 2016 Aug 15;11(8):e0160460. [doi: 10.1371/journal.pone.0160460] [Medline: 27525413]
2. Petersen M, Yiannoutsos CT, Justice A, Egger M. Observational research on NCDs in HIV-positive populations: conceptual and methodological considerations. J Acquir Immune Defic Syndr 2014 Sep 01;67 Suppl 1(0 1):S8-S16 [FREE Full text] [doi: 10.1097/QAI.0000000000000253] [Medline: 25117964]
3. Poorolajal J, Hooshmand E, Mahjub H, Esmailnasab N, Jenabi E. Survival rate of AIDS disease and mortality in HIV-infected patients: a meta-analysis. Public Health 2016 Oct;139:3-12. [doi: 10.1016/j.puhe.2016.05.004] [Medline: 27349729]
4. Hoffmann CJ, Fielding KL, Johnston V, Charalambous S, Innes C, Moore RD, et al. Changing predictors of mortality over time from cART start: implications for care. J Acquir Immune Defic Syndr 2011 Nov 01;58(3):269-276 [FREE Full text] [doi: 10.1097/QAI.0b013e31823219d1] [Medline: 21876447]
5. Grinsztejn B, Luz PM, Pacheco AG, Santos DVG, Velasque L, Moreira RI, et al. Changing mortality profile among HIV-infected patients in Rio de Janeiro, Brazil: shifting from AIDS to non-AIDS related conditions in the HAART era. PLoS One 2013;8(4):e59768. [doi: 10.1371/journal.pone.0059768] [Medline: 23577074]
6. Triant VA. Cardiovascular disease and HIV infection. Curr HIV/AIDS Rep 2013 Sep;10(3):199-206 [FREE Full text] [doi: 10.1007/s11904-013-0168-6] [Medline: 23793823]
7. Kazooba P, Kasamba I, Mayanja BN, Lutaakome J, Namakoola I, Salome T, et al. Cardiometabolic risk among HIV-POSITIVE Ugandan adults: prevalence, predictors and effect of long-term antiretroviral therapy. Pan Afr Med J 2017 May 15;27:40 [FREE Full text] [doi: 10.11604/pamj.2017.27.40.9840] [Medline: 28761616]
8. Kiage JN, Heimburger DC, Nyirenda CK, Wellons MF, Bagchi S, Chi BH, et al. Cardiometabolic risk factors among HIV patients on antiretroviral therapy. Lipids Health Dis 2013 Apr 10;12:50 [FREE Full text] [doi: 10.1186/1476-511X-12-50] [Medline: 23575345]
9. Sogarwa R, Mehra S. Approaches to address NCD among PLHIV in low and middle income counties. J AIDS Clin Res 2015;06(06):1000472 [FREE Full text] [doi: 10.4172/2155-6113.1000472]
10. Nigatu T, Oldenburg B, Elliott J, Setswe G, Woldegiorgis MA. The incidence of cardiovascular disease, cancer and type 2 diabetes comorbidities in HIV infection: A systematic review. J Nurs Educ Pract 2013 Jan 09;3(7):58-67. [doi: 10.5430/jnep.v3n7p58]
11. De Wit S, Sabin CA, Weber R, Worm SW, Reiss P, Cazanave C, Data Collection on Adverse Events of Anti-HIV Drugs (D:A:D) study. Incidence and risk factors for new-onset diabetes in HIV-infected patients: the Data Collection on Adverse Events of Anti-HIV Drugs (D:A:D) study. Diabetes Care 2008 Jun;31(6):1224-1229 [FREE Full text] [doi: 10.2337/dc07-2013] [Medline: 18268071]
12. Haregu TN, Oldenburg B, Setswe G, Elliott J. Magnitude of diabetes comorbidity among people living with HIV: A systematic review. Int J Diabetes Res 2012 Dec 7;1(5):81-86 [FREE Full text] [doi: 10.5923/j.diabetes.20120105.02]
13. Brown TT, Cole SR, Li X, Kingsley LA, Palella FJ, Riddler SA, et al. Antiretroviral therapy and the prevalence and incidence of diabetes mellitus in the multicenter AIDS cohort study. Arch Intern Med 2005 May 23;165(10):1179-1184. [doi: 10.1001/archinte.165.10.1179] [Medline: 15911733]
14. Chow FC, Regan S, Feske S, Meigs JB, Grinspoon SK, Triant VA. Comparison of ischemic stroke incidence in HIV-infected and non-HIV-infected patients in a US health care system. J Acquir Immune Defic Syndr 2012 Aug 01;60(4):351-358 [FREE Full text] [doi: 10.1097/QAI.0b013e31825c7f24] [Medline: 22580566]
15. Tseng ZH, Secemsky EA, Dowdy D, Vittinghoff E, Moyers B, Wong JK, et al. Sudden cardiac death in patients with human immunodeficiency virus infection. J Am Coll Cardiol 2012 May 22;59(21):1891-1896 [FREE Full text] [doi: 10.1016/j.jacc.2012.02.024] [Medline: 22595409]
16. Butt AA, Chang C, Kuller L, Goetz MB, Leaf D, Rimland D, et al. Risk of heart failure with human immunodeficiency virus in the absence of prior diagnosis of coronary heart disease. Arch Intern Med 2011 May 25;171(8):737-743 [FREE Full text] [doi: 10.1001/archinternmed.2011.151] [Medline: 21518940]
17. Currier JS, Taylor A, Boyd F, Dezii CM, Kawabata H, Burtcel B, et al. Coronary heart disease in HIV-infected individuals. J Acquir Immune Defic Syndr 2003 Aug 01;33(4):506-512. [doi: 10.1097/00126334-200308010-00012] [Medline: 12869840]
18. Obel N, Thomsen HF, Kronborg G, Larsen CS, Hildebrandt PR, Sørensen HT, et al. Ischemic heart disease in HIV-infected and HIV-uninfected individuals: a population-based cohort study. Clin Infect Dis 2007 Jul 15;44(12):1625-1631. [doi: 10.1086/518285] [Medline: 17516408]
19. Triant VA, Lee H, Hadigan C, Grinspoon SK. Increased acute myocardial infarction rates and cardiovascular risk factors among patients with human immunodeficiency virus disease. J Clin Endocrinol Metab 2007 Jul;92(7):2506-2512 [FREE Full text] [doi: 10.1210/jc.2006-2190] [Medline: 17456578]

20. Mdege ND, Shah S, Ayo-Yusuf OA, Hakim J, Siddiqi K. Tobacco use among people living with HIV: analysis of data from Demographic and Health Surveys from 28 low-income and middle-income countries. Lancet Glob Health 2017 Jun;5(6):e578-e592 [FREE Full text] [doi: 10.1016/S2214-109X(17)30170-5] [Medline: 28495263]
21. Saitz R. Clinical practice. Unhealthy alcohol use. N Engl J Med 2005 Feb 10;352(6):596-607. [doi: 10.1056/NEJMcp042262] [Medline: 15703424]
22. Lasica JD. The Mobile Generation: Global Transformations At the Cellular Level: A report of the Fifteenth annual Aspen Institute Roundtable on Information Technology. Washington, DC: Aspen Institute Communications and Society Program; 2007.
23. Catalani C, Philbrick W, Fraser H, Mechael P, Israelski DM. mHealth for HIV Treatment & Prevention: A Systematic Review of the Literature. Open AIDS J 2013 Aug 13;7:17-41 [FREE Full text] [doi: 10.2174/1874613620130812003] [Medline: 24133558]
24. Hirsch-Moverman Y, Daftary A, Yuengling KA, Saito S, Ntoane M, Frederix K, et al. Using mHealth for HIV/TB treatment support in Lesotho: Enhancing patient-provider communication in the START Study. J Acquir Immune Defic Syndr 2017 Jan 01;74 Suppl 1(Suppl 1):S37-S43 [FREE Full text] [doi: 10.1097/QAI.0000000000001202] [Medline: 27930610]
25. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen PA, Li YCJ, et al. mHealth: An updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. Comput Methods Programs Biomed 2015 Nov;122(2):257-265. [doi: 10.1016/j.cmpb.2015.08.003] [Medline: 26304621]
26. Tufts KA, Johnson KF, Shepherd JG, Lee JY, Ajzoon MSB, Mahan LB, et al. Novel interventions for HIV self-management in African American women: a systematic review of mHealth interventions. J Assoc Nurses AIDS Care 2015;26(2):139-150. [doi: 10.1016/j.jana.2014.08.002] [Medline: 25283352]
27. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012 Mar 14;3(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
28. Park LG, Howie-Esquivel J, Dracup K. A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence. J Adv Nurs 2014 Sep;70(9):1932-1953. [doi: 10.1111/jan.12400] [Medline: 24689978]
29. van Velthoven MHMMT, Brusamento S, Majeed A, Car J. Scope and effectiveness of mobile phone messaging for HIV/AIDS care: a systematic review. Psychol Health Med 2013;18(2):182-202. [doi: 10.1080/13548506.2012.701310] [Medline: 22788357]
30. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. BMC Health Serv Res 2015 Jan 22;15(1):33 [FREE Full text] [doi: 10.1186/s12913-014-0654-6] [Medline: 25609559]
31. PRISMA. PRISMA Transparent Reporting of Systematic Reviews and Meta-Analyses. URL: http://www.prisma-statement.org/ [accessed 2021-05-27]
32. mHealth: New horizons for health through mobile technologies. World Health Organization. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2021-05-27]
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
34. Module 4: Selecting studies and collecting data. The Cochrane Collaboration. URL: http://training.cochrane.org/resource/data-collection-forms-intervention-reviews [accessed 2021-05-27]
35. Anglada-Martínez H, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Sequeira E, Aragunde V, et al. Feasibility and preliminary outcomes of a web and smartphone-based medication self-management platform for chronically ill patients. J Med Syst 2016 Apr;40(4):99. [doi: 10.1007/s10916-016-0456-y] [Medline: 26872781]
36. Morillo-Verdugo R, Robustillo-Cortés MDLA, Martín-Conde MT, Callejón-Callejón G, Cid-Silva P, Moriel-Sánchez C, et al. Effect of a structured pharmaceutical care intervention versus usual care on cardiovascular risk in HIV patients on antiretroviral therapy: INFAMERICA Study. Ann Pharmacother 2018 Nov;52(11):1098-1108. [doi: 10.1177/1060028018778045] [Medline: 29808711]
37. Roos R, Myezwa H, van Aswegen H, Musenge E. Effects of an education and home-based pedometer walking program on ischemic heart disease risk factors in people infected with HIV: a randomized trial. J Acquir Immune Defic Syndr 2014 Nov 01;67(3):268-276. [doi: 10.1097/QAI.0000000000000299] [Medline: 25140908]
38. Zuñiga J, García AA, Silva L, Park JM, Barrera Y. Pilot test of an educational intervention to improve self-management of diabetes in persons living with HIV. Pilot Feasibility Stud 2019 Sep 07;5:111 [FREE Full text] [doi: 10.1186/s40814-019-0495-5] [Medline: 31516728]
39. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
40. Grinspoon S. Study of lifestyle modification in HIV lipodystrophy. National Institutes of Health, U.S. National Library of Medicine, ClinicalTrials.gov. 2006. URL: https://clinicaltrials.gov/ct2/show/NCT00111358 [accessed 2021-05-27]

41. Jaggers JR, Dudgeon W, Blair SN, Sui X, Burgess S, Wilcox S, et al. A home-based exercise intervention to increase physical activity among people living with HIV: study design of a randomized clinical trial. BMC Public Health 2013 May 24;13(1):502 [FREE Full text] [doi: 10.1186/1471-2458-13-502] [Medline: 23706094]
42. Brooke H. iSTEP - an mHealth physical activity and diet intervention for persons with HIV. National Institutes of Health, U.S. National Library of Medicine, ClinicalTrials.gov. 2017. URL: https://clinicaltrials.gov/ct2/show/NCT03123731 [accessed 2021-05-28]
43. Dodson S, Klassen KM, McDonald K, Millard T, Osborne RH, Battersby MW, et al. HealthMap: a cluster randomised trial of interactive health plans and self-management support to prevent coronary heart disease in people with HIV. BMC Infect Dis 2016 Mar 05;16(1):114 [FREE Full text] [doi: 10.1186/s12879-016-1422-5] [Medline: 26945746]
44. Oduor E, Nyota T, Wachira C, Osebe S, Remy SL, Walcott A. Medication management companion (MMC) for a rural Kenyan community. In: CSCW '18: Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2018 Oct 30 Presented at: CSCW '18: Computer Supported Cooperative Work and Social Computing; November 3-7, 2018; Jersey City, NJ, USA p. 145-148. [doi: 10.1145/3272973.3274041]
45. Kengne AP. Mobile phone SMS for adherence support for hypertension in South African adults in care for HIV infection. Pan African Clinical Trials Registry. 2019. URL: https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=4711 [accessed 2021-05-28]
46. 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] [Medline: 25803266]
47. Deaths. World Health Organization. URL: https://www.who.int/gho/hiv/en/ [accessed 2020-07-07]

Abbreviations

ART: antiretroviral therapy
CMD: cardiometabolic disease
LMIC: low- and middle-income country
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
RCT: randomized controlled trial
WHO: World Health Organization

Edited by L Buis; submitted 16.05.20; peer-reviewed by J Seitz, C Wu; comments to author 11.07.20; revised version received 13.08.20; accepted 16.04.21; published 09.06.21.

Please cite as:
Odukoya OO, Ohazurike C, Akanbi M, O'Dwyer LC, Isikekpei B, Kuteyi E, Ameh IO, Osadiaye O, Adebayo K, Usinoma A, Adewole A, Odunukwe N, Okuyemi K, Kengne AP
mHealth Interventions for Treatment Adherence and Outcomes of Care for Cardiometabolic Disease Among Adults Living With HIV: Systematic Review
JMIR Mhealth Uhealth 2021;9(6):e20330
URL: https://mhealth.jmir.org/2021/6/e20330
doi:10.2196/20330
PMID:34106075

Original Paper

Acceptability of a Mobile Phone Support Tool (Call for Life Uganda) for Promoting Adherence to Antiretroviral Therapy Among Young Adults in a Randomized Controlled Trial: Exploratory Qualitative Study

Adelline Twimukye[1], BA, MSc; Agnes Bwanika Naggirinya[1], MD, MSc, MMed; Rosalind Parkes-Ratanshi[1,2], MBBS, MA, PhD; Ronnie Kasirye[1], MD, MSc, PhD; Agnes Kiragga[1], BStat, BCom, MSc, PhD; Barbara Castelnuovo[1], MD, MMed, PhD; Jacob Wasswa[1], BIT, BSc; Maria Sarah Nabaggala[1], BStat, BCom, MSc; Elly Katabira[1,3], MD, MMed, PhD; Mohammed Lamorde[1], MD, MMed, PhD; Rachel Lisa King[4], DiP, BA, MPH, PhD

[1]Infectious Diseases Institute, College of Health Sciences, Makerere University, Kampala, Uganda

[2]Department of Public Health & Primary Care, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom

[3]College of Health Sciences, School of Medicine, Makerere University, Kampala, Uganda

[4]Institute for Global Health Sciences, University of California, San Francisco, CA, United States

Corresponding Author:
Agnes Bwanika Naggirinya, MD, MSc, MMed
Infectious Diseases Institute
College of Health Sciences
Makerere University
P.O. Box 22418
University Hall Lane
Kampala, 10218
Uganda
Phone: 256 256 312 307000
Email: anaggirinya@idi.co.ug

Abstract

Background: Adherence to treatment is critical for successful treatment outcomes. Although factors influencing antiretroviral therapy (ART) adherence vary, young adults are less likely to adhere owing to psychosocial issues such as stigma, ART-related side effects, and a lack of access to treatment. The Call for Life Uganda (CFLU) mobile health (mHealth) tool is a mobile phone–based technology that provides text messages or interactive voice response functionalities through a web interface and offers 4 modules of support.

Objective: This study aims to describe the acceptability and feasibility of a mobile phone support tool to promote adherence to ART among young adults in a randomized controlled trial.

Methods: An exploratory qualitative design with a phenomenological approach at 2 study sites was used. A total of 17 purposively selected young adults with HIV infection who had used the mHealth tool CFLU from 2 clinics were included. In total, 11 in-depth interviews and 1 focus group discussion were conducted to examine the following topics: experience with the CFLU tool (benefits and challenges), components of the tool, the efficiency of the system (level of comfort, ease, or difficulty in using the system), how CFLU resolved adherence challenges, and suggestions to improve CFLU. Participants belonged to 4 categories of interest: young adults on ART for the prevention of mother-to-child transmission, young adults switching to or on the second-line ART, positive partners in an HIV-discordant relationship, and young adults initiating the first-line ART. All young adults had 12 months of daily experience using the tool. Data were analyzed using NVivo version 11 software (QSR International Limited) based on a thematic approach.

Results: The CFLU mHealth tool was perceived as an acceptable intervention; young adults reported improvement in medication adherence, strengthened clinician-patient relationships, and increased health knowledge from health tips. Appointment reminders and symptom reporting were singled out as beneficial and helped to address the problems of forgetfulness and stigma-related issues. HIV-related stigma was reported by a few young people. Participants requested extra support for scaling up CFLU to

make it more youth friendly. Improving the tool to reduce technical issues, including network outages and a period of software failure, was suggested. They suggested that in addition to digital solutions, other support, including the promotion of peer support meetings and the establishment of a designated space and staff members for youth, was also important.

Conclusions: This mHealth tool was an acceptable and feasible strategy for improving ART adherence and retention among young adults in resource-limited settings.

Trial Registration: ClinicalTrials.gov NCT02953080; https://clinicaltrials.gov/ct2/show/NCT02953080

(JMIR Mhealth Uhealth 2021;9(6):e17418) doi:10.2196/17418

KEYWORDS

HIV; mHealth; young adults; adherence; qualitative; Uganda

Introduction

Background

In 2018, of the 3.9 million young people living with HIV globally, 80% were in Sub-Saharan Africa [1]. The 2017 Uganda Bureau of Statistics report indicated that more than half of Uganda's population (55%) is aged under 18 years [2]. The burden of HIV among youth in Uganda is high. In Uganda, HIV prevalence triples from those aged 15-19 years (1.1%) to those aged 20-24 years (3.3%), suggesting that new infections remain an issue in this age group [3]. Among young adults (aged 15-24 years), there is a disparity in HIV prevalence by sex. HIV prevalence is almost 4 times higher among women than men aged 15-19 years and 20-24 years [4].

HIV care and treatment services in the region should adapt to adequately meet the antiretroviral therapy (ART) demand to suppress the viral load to undetectable levels and improve health outcomes. Suboptimal ART adherence over time predicts virological failure, development of drug resistance, and death [5-8]. Adolescents and young adults have poorer rates of adherence to ART and retention in care than older adults [9]. Young adults are less likely to adhere to ART medication if they have psychosocial issues such as stigma, previously experienced side effects with the treatment, and lack of access to treatment [10-12].

In 2018, 50% of the population (24.7 million) were phone subscribers (fixed or cellular) [13]. Compared with adults (40%), young adults (66%) are more likely to use mobile phone SMS text messaging [13]. Mobile phones have become a common accessory for most young people in Uganda. Self-reported quantitative survey data collected in 2008-2009 from 1503 secondary school students in Mbarara, Uganda, suggest that 27% currently had mobile phones and about half (51%) of all students and 61% of those who owned a mobile phone believed that they would access an SMS text messaging–based HIV prevention program [14]. One study assessed the patterns and dynamics of mobile phone use among an ART cohort in rural Uganda, ascertained its feasibility for improving clinic attendance, and found that mobile phones have the potential for use in resource-constrained settings to improve the clinical management of people living with HIV [15]. Texting-based and other mobile health (mHealth) interventions could be particularly suitable for young adults who, irrespective of their socioeconomic status, are typically competent users of mobile phones and text messages. Globally, mHealth interventions have resulted in improvements in treatment adherence for patients with asthma, diabetes, and tuberculosis [16].

Objectives

Various mHealth studies have shown improved virological outcomes [17,18], lower risk of nonadherence [19-21], and good retention in care before ART initiation [15]. However, there is little specific evidence on the effect of mHealth on young adults living with HIV. We used qualitative methods to describe the acceptability of using Call for Life Uganda (CFLU, an adherence support tool offering interactive voice response [IVR] or SMS) to promote adherence to ART for young adults in a randomized controlled trial (RCT).

Methods

Study Design

We implemented an exploratory qualitative design using a phenomenological approach to describe the acceptability and feasibility of a mobile phone support tool to promote adherence to ART among young adults. The study was nested within an open-label RCT titled *Improving Outcomes in HIV Patients Using Mobile Phone–Based Interactive Software Support*. The technology evaluated in this study is connect for life (CfL), a MOTECH-based, open-source platform developed by the Grameen Foundation and the University of Southern Maine. It was supported by Janssen, the Pharmaceutical Companies of Johnson and Johnson, and was released under the terms of MOTECH's open-source license agreement. We adapted CfL for use among people living with HIV and named the system CFLU. CFLU allows a computer to interact with patients using voice and tone input via keypad (IVR) or by SMS, with the choice made by the patient. All content was developed in English and 2 additional local languages, following the Uganda HIV treatment guidelines. People living with HIV in the control arm received standard of care comprising face-to-face facility appointments, no remote adherence or appointment reminders, and no remote symptom reporting. In the intervention arm, the people living with HIV received standard of care and daily adherence IVR call or SMS text messaging at the preplanned times of taking ART. In the intervention arm, the people living with HIV also received a dedicated call offering a health education message (a health tip) weekly. Patients received IVR calls or SMS text message appointment reminders on or before the scheduled appointment date. The CFLU platform allows patients to report symptoms at the end of the scheduled call, or at any time, through a toll-free line. Multimedia Appendix 1

illustrates the call flow from the CFLU tool to the end user. Participants chose the preferred languages, time, and frequency of receiving calls.

In this qualitative study, we used in-depth interviews (IDIs) and 1 focus group discussion (FGD) to gain insights into individual and community perspectives about the CFLU tool. The purpose of the FGD in this study was to stimulate group conversations on specific themes, to assess the differences and similarities in perceptions, values, norms, and preferences among young adults. We explored only the normative and topics that were not considered sensitive in the FGD [22].

Study Participants and Selection

Participants were enrolled from 2 HIV clinics acting as study sites for the CFLU RCT (the Infectious Diseases Institute clinic at Mulago National Referral Hospital, Kampala, and Kasangati Health Centre facility level IV, a periurban government health facility). Purposive sampling of young adults with HIV infection aged 18-25 years was carried out from these 2 clinics.

The total number of young adults was 161. We undertook purposive sampling within this group to recruit them for this study. Inclusion criteria were registration in the CfL RCT, had access to a mobile phone, had the ability to use basic mobile phone functions such as making and receiving calls, and willingness to comply with scheduled visits. Patients who understood 1 of the 2 local languages or English were eligible for the study. Young adults who were critically ill; those aged under 18 years; and those whose clinical condition interfered with the appropriate use of their mobile phone, such as deafness or severe cognitive impairment, were excluded from the study. The young adults could also belong to other groups of interest enrolled in the study, including pregnant women receiving ART for the prevention of mother-to-child transmission, those switching to or on second-line ART, those initiating first-line ART, and positive partners in a discordant relationship.

Data Collection

We made telephone calls to potential young adults to brief them about the study and invite them for the study. We negotiated an appropriate time, place, and day for the interview with the IDI and FGD participants and booked secure private locations to conduct interviews at the Infectious Diseases Institute clinic and Kasangati, as appropriate. Clinic study staff, a counselor with social science expertise and Good Clinical Practice, and a postdoctoral scholar trained in FGD note taking collected the data. The FGD had a range of 6-12 participants. The clinic study staff contacted potential participants by phone, described the study, and extended an invitation to participate. We conducted 2 interviews in English and 10 in the local language (Luganda) at the Infectious Diseases Institute and Kasangati in private locations (IDIs) and community centers (FGDs). Before data collection, we translated the topic guides from English into Luganda, the main language spoken in the catchment area of the study clinics. Each IDI and FGD lasted for approximately 1 hour and were audio recorded and complemented by written notes. We examined the following topics in IDIs and FGD: (1) experience with the CFLU tool (likes and dislikes), (2) components of the tool (health tips, pill and appointment reminders, and symptom reporting), (3) efficiency and willingness to pay for the system (level of comfort, ease, or difficulty in using the system), (4) how CFLU resolved adherence challenges, and (5) suggestions to improve CFLU. All young adults had 12 months of daily experience using the tool. Young adults were recruited and interviewed until no new themes emerged, and saturation was reached when there was repetition of the previously mentioned themes.

Data Analysis and Interpretation

All audio recordings from IDIs and FGDs were transcribed verbatim after translation by independent professionals. The research team read the typed transcripts several times and filled in the gaps by listening to the audio recordings. Data were managed and analyzed using NVivo version 11 (QSR International Limited). Each independent transcript was read and reread by a senior social scientist for emergent themes and recurrent ideas and then aggregated into themes. Codes were assigned to relevant segments of the text; similar codes were aggregated to form themes, which were then used to address the research questions and develop coherent narratives. The team developed an explicit codebook describing each category and theme. In the next step, the team sorted the quotes based on themes. The team then examined the degree to which these themes were distributed across gender, age group, and social target group. After team members read the interview transcripts, data were coded into meaning units and major themes were developed. Quotations and key phrases are highlighted in the findings.

Ethical and Regulatory Approval

The CFLU RCT was approved by the Makerere University School of Public Health Higher Degrees Research and Ethics and Committee (Number 378) and the Uganda National Council of Science and Technology (Number Health Sciences 3005) and was registered with ClinicalTrials.gov (NCT02953080). All the study procedures, compensation, benefits, potential risk of participation, and voluntary and confidential nature of participation were discussed. Written informed consent was obtained from all respondents before enrollment in the qualitative study. For young adults with low literacy, we used a thumbprint in the presence of a witness.

Results

Overview

A total of 82 young adults who used the CFLU tool were given a telephone call, inviting them for FGD or IDIs (Multimedia Appendix 2). Of the 82 young adults, 37 were found ineligible as their age limit was slightly more than that required for young adults (26-28 years) and 28 could not make it for the discussion because of various reasons such as busy schedules and phones not available owing to poor network. A total of 21 young adults agreed to participate, 17 on the intervention arm and 4 on the control arm (analyzed separately and not included here (Multimedia Appendix 2).

A total of 11 IDIs and 1 FGD with those on the intervention arm were conducted. A total of 6 young adults were from the Infectious Diseases Clinic, whereas 11 were from Kasangati

Health Center IV. The age range of young adults was 18-25 years; the majority were females, and more than 81% (17/21) of them were in a steady sexual relationship. Most young adults had a formal education, with only 14% (3/21) having reached the tertiary level (Table 1).

The results covered 4 major themes related to CFLU. These were attitudes toward CFLU, components of CFLU, barriers to and challenges of CFLU, and suggestions or recommendations from youth to improve CFLU. The following sections organize the results into 3 levels: personal and household factors, facility or service delivery factors, and community factors (Table 2).

Table 1. Demographics of participants who took part in the qualitative study (N=21)[a].

Population demographics	Participants, n (%)
Age (years)	
18-19	3 (14)
20-24	18 (85)
Gender	
Male	5 (24)
Female	16 (76)
Marital status	
In a relationship	17 (81)
Single	4 (19)
Education status	
Secondary education and primary education	15 (71)
Tertiary education	3 (14)
No formal education	1 (5)
Missing information	2 (10)

[a]Four young adults were excluded because they had no experience with the tool.

Table 2. Frequency of each theme for young adults on CFLU[a] (N=21).

Theme and categories	Patients, n (%)	Number of responses
Positive perceptions toward the CFLU[a] tool		
CFLU system perceived as an adherence supporter	6 (29)	18
Management of HIV stigma–related challenges	6 (29)	18
Increased psychosocial support	10 (48)	25
Strengthened clinician-patient relationship	5 (24)	7
The kind tone of voice of the system	3 (14)	6
Confidentiality and privacy	9 (43)	13
Ease of system use	7 (33)	11
Components of the CFLU tool		
Clinic visit call reminders	12 (57)	34
Symptoms reporting	12 (57)	46
Health tips call	12 (57)	44
Daily pill reminder call	12 (57)	57
Barriers to and challenges of the CFLU tool		
Technical issues with the system	11 (52)	22
Poor access to mobile phones	12 (57)	23
Fear of HIV-related stigma	6 (29)	18
Suggestions to improve the CFLU tool		
Resolving adherence challenges	5 (24)	9
Suggesting additional health tips	12 (57)	45
Providing adolescent peer support	7 (33)	20
Scaling up CFLU	8 (38)	11
Sensitizing about CFLU and HIV	9 (43)	16
Making CFLU appealing	2 (10)	8
Getting the designated space for CFLU	1 (5)	3
Getting the designated staff for CFLU	5 (24)	10
Resolving the stigma	4 (19)	6
Combining the CFLU appointment with routine clinic appointment	31 (4)	5
Other suggestions	4 (19)	4

[a]CFLU: Call for Life Uganda.

Positive Perceptions Toward CFLU

Personal and Household Related

CFLU System Perceived as an Adherence Supporter

Most young adults from both sites, but mainly from Kasangati, perceived the CFLU system as an aid to adherence through its pill reminders, especially for youth who had trouble taking their medication and who forgot to take their medication because of busy work and school schedules.

> *It gives me strength, when I had just started my treatment I was very scared. So, I didn't know...how to take my medicine....The pill was very large...but the way the answering machine speaks, it gives me courage to really take my pills*. [FGD, female, 18-24 years, intervention, Kasangati]

CFLU comes in as a treatment supporter to remind one to take medication; it is like a parent asking, "Hey, did you swallow your medication?" It is a voice message but sounds like a person who is aware of one's HIV status.

Resolving Forgetfulness

Both male and female youth across different groups admitted to poor adherence to ART in the past and believed that CFLU would resolve most of the barriers to adherence, such as forgetfulness. The reminder call was deemed critical when a young adult slept off or during a busy school or work schedule:

> *I used to forget to take pills on time. But the doctor cares so much and calls to remind me to take my medication...majority miss taking drugs but each time a call comes through, it reminds you, most youth get tired mentally; they forget.* [IDI, 18 years, male, intervention, Kasangati]
>
> *CFLU is good because it helps me take my pills on time, whenever I am very tired and fatigued; I sleep off before the time for taking pills... when my phone rings, I wake up immediately, and straight away I take my drugs.* [FGD, female, 18-24 years, intervention Kasangati]

There was a perception among most young adults from both sites that the tool improved their health through viral load suppression. Most youth on CFLU said that pill reminders helped them take drugs properly, daily, and on time, which led to viral suppression compared with when they were not on the system:

> *It's helpful to improve pill adherence & keeping appointments [that] will cause improvement of health, if the viral load is reducing then the opportunistic infections are reducing.* [IDI, female, 24 years, intervention arm, Mulago]

Management of HIV Stigma–Related Challenges

Most young adults, mainly from Infectious Diseases Institute clinic, stated that CFLU resolved the fear of stigma through its health tips that continuously educated them about how to overcome stigma and emphasized the importance of disclosure to avoid stigma and the importance of adherence to ART. Adherence to ART was easier in the context of young adults on CFLU who had not disclosed their HIV status:

> *Call for life does continuous education about overcoming stigma. It advises and teaches us to overcome stigma which encourages us in our HIV situation, we (youth) cease to be fearful.* [IDI, 24 years, intervention, Mulago]
>
> *We follow positive living health tips, they tell us "you don't have to worry about your status when you are HIV positive" you know we are always worried, some of us have stigma. They just teach us to stay and live positive, eat a balance diet.* [FGD, female, 18-24 years, intervention arm, Kasangati]
>
> *The fact that we are not supposed to talk back after receiving the call, sometimes when we are in public places, we do not have to worry that someone will hear what you are saying you just listen and keep following the prompts and others will not know anything. It works for me personally.* [FGD, female, 18-24 years, intervention arm, Kasangati]

Increased Psychosocial Support

Most young adults, mainly from Kasangati, expressed that they received psychosocial support from the tool, especially in those starting ART. Emotionally, the youth on CFLU felt that they were not alone. Through interaction with the loving voice that instructed them on how to take treatment, they were able to obtain coping support and gain hope for the future. The counseling provided during CFLU reduced high-risk behavior and promoted their own psychological well-being:

> *There is a way you do not lose hope when you are HIV infected on CFLU....Personally, I have liked it because it encourages, counsels and one does not lose hope.* [IDI, female, 22 years, intervention arm, Kasangati]
>
> *Before I had multiple sexual partners but now ever since I was taught about risk reduction through health tips, I realized that I was wasting my life then I stopped....* [IDI, female, 23 years, intervention arm, Mulago]

Facility or Service Delivery Factors

Strengthened the Clinician-Patient Relationship

The young adults, mainly from IDI, expressed that the tool improved the relationship between health care workers and patients, making the patients feel cared for. The youth said that there was a good relationship and interaction between patients and CFLU doctors. Paying attention to detail, politeness, care, and treatment, as well as offering transport refund by doctors enhanced proper treatment adherence that would lead to health improvement for youth:

> *There is that element of a good relationship. You see the moment you get used to someone, you try to open up about your challenges; there is that bridging of the gap, that closeness that is created between the two parties.* [IDI, male, 24 years, intervention arm, Mulago]

The Kind Tone of Voice of the System

Most participants from both sites perceived that the tone of voice and the wording used through the system was kind and conveyed friendliness, trust, and care by the health providers. They looked forward to the call that reminded them to take their drugs properly on time:

> *It's so soothing and comforting and it makes you feel good. The man says, hallo my beloved, by this time you should have taken your medicine, for sure, you really feel loved and cared for. I really like that voice system.* [FGD, female, 18-24 years, intervention, Kasangati]
>
> *The one speaking speaks in a calm and thankful manner, says "beloved thank you for listening," and the voice is good.* [IDI, female, 23 years, intervention arm, Mulago]

Most youth liked CFLU because it taught them how to swallow medicine on time and how to live positively through the routine health tips.

Community Factors

Confidentiality and Privacy

Most participants from both sites, but mainly from Kasangati, liked confidentiality and privacy through the use of the individualized secret pin code. Young adults were taught how

to use the system, especially how to enter their secret pin number, and they found it easy to use:

> *I was taught, when they call, I have to enter my pin number. Then that is when they will speak. If you don't put in the pin it won't talk.* [IDI, female, 23 years, intervention arm, Mulago]

Reducing the volume of the phone as they listened to voice communication maintained their HIV status privacy, as one participant said:

> *I reduce the volume of my phone, even if someone notices me answering the call, they can't know what I'm up to, and they may even ask, eh, you have received a call, how come you are not talking?* [IDI, female, 19 years, intervention arm, Mulago]

Ease of Use for the System

The majority of young adults from both sites found the system easy to use and attributed this to the training or orientation they had received. The system also allows convenient settings for young adults, including individualized pill reminders and time to receive health tips:

> *They asked me to give them a time that is convenient for me and I choose 9:05 o'clock, and when that time comes, the phone rings to remind me to swallow my medicine.* [IDI, female, 19 years, intervention arm, Mulago]

Components of CFLU

Beneficial Clinic Visit Call Reminders

All young adults described CFLU as a call system that emphasized mainly pill reminders, health information tips, clinic appointment reminders, and symptom self-reporting, and they were all useful as far as adherence support was concerned.

Most young adults, but mainly from IDI, said that CFLU enhanced keeping routine clinic appointments that promoted ART adherence. Clinic appointment reminders were considered beneficial for those who misplaced cards or forgot appointments. Clinic visit remainders enabled young adults to plan for payment for transport:

> *When in a faraway place, so when you get a reminder 2 days before the appointment, you plan for transport.* [FGD, female, 18-24 years, intervention arm, Kasangati]

> *Appointment reminders are going to be helpful because I have ever missed an appointment, by the time I checked the card the time had passed, so it's helpful if am reminded.* [IDI, male, 23 years, intervention arm, Mulago]

Value of Symptom Reporting

Through symptom reporting, young adults from both sites, but mainly from IDI, were able to seek treatment and medical advice for themselves or their child. Young adults reported symptoms such as headache, diarrhea, fever, chest pain, abdominal pain, cough, and constipation:

> *I saw its advantage when I fell sick...CFLU is responsible, when you report your symptom such as cough or fever. You get better or advice and you are told to seek medical attention if the symptom worsens....* [FGD, female, young adults, 18-24 years, intervention arm, Kasangati]

> *I had a running stomach one time and a fever...Then I pressed the option for one with a symptom. The doctor called me and told me to come to the clinic the next day and they diagnosed me.* [IDI, female, 19 years, intervention arm, Mulago]

However, delayed responses from health care workers were a problem for most young adults after self-reporting a symptom. By the time the doctor called back, the symptoms often had disappeared. They recommended that the doctor should call back within 20-30 minutes:

> *I had a headache, I reported in the night and the doctor called back in the morning but when she called I had swallowed some pain killers.* [FGD, female, 18-25 years, intervention arm, Kasangati]

Use of Health Tips Call

Access to information through health tips increased patient knowledge in young adults from both sites, but mainly IDI. The most frequent health tips selected were positive living, HIV, nutrition, avoidance of risky sexual behaviors such as having multiple partners, alcohol and drug abuse, breastfeeding options, and prevention of mother-to-child transmission. Few young adults reported the fear of stigma and discrimination that was addressed in the health tips messaging:

> *Initially, at home, it was worse; I suffered low self-esteem due to HIV, as a result, I used to keep everything to myself because I did not want my family members to see me take HIV medication and point fingers at me. I feel I am now a free person after listening to the health tips on avoiding stigma. Call for life reminds me to take my drugs without any body noticing.* [IDI, male, 19 years, intervention arm, Mulago]

The belief that antiretroviral therapy (ART) cause undesirable side effects that may affect drug adherence was reported by most youth at IDI. Through health tips, young people received basic ART education:

> *When you have just started drugs let's say within a month ...one may get dizziness and begin noticing changes..., when they educate us before taking drugs they tell us, expect side effects. You may expect weird dreams, dizziness, that's why they say you have to choose your appropriate time.* [IDI, female, 22 years, intervention arm, Mulago]

Majority of young adults attributed good health outcomes to these health tips. For example, one young woman who recently delivered an HIV-negative baby said:

> *If you take drugs properly without missing, you reduce the chances of infecting the baby. I personally gave*

> *birth to a negative baby; I took drugs from here.* [IDI, female, 22 years, intervention arm, Kasangati]

However, few young adults felt there could be more variation in the health tips offered or that the tips were not relevant to them:

> *They should not give the same health tips every day, they should ask one, if you want to listen to tips about relationships, you press 1, if you want about pregnancy, then you press that option.* [IDI, female, 22 years, intervention arm, Mulago]

Majority of young adults requested more diverse and additional health information. Positive female partners in discordant relationships requested for more health information about healthy relationships, disclosure in sero-discordancy, and conception in sero-discordant relationships. Textbox 1 summarizes the different requests.

Textbox 1. Health tips requested by young adults to be added to the Call for Life Uganda system.

Health tips and illustrative quotes

- HIV discordancy and relationship advice
  - "Let call for life explain to us more about Prophylactic use of ARVS in a discordant relationship. If one is in a discordant relationship and has been on ARVs for a while, when viral load is fine, the doctors still insist on condom use and you are like for sure! For how long?" (Focus group discussion [FGD], female, 18-24 years, intervention, Kasangati)
- Female condom use
  - "I want CFL to add tips on how to use female condoms...I totally failed to use them." (FGD, female, 18-24 years, intervention arm, Kasangati)
- Relationship handling
  - "Actually, talk about the relationships. I think C4L can come in and educate when to have a partner and how to identify and choose a rightful partner." (in-depth interview [IDI], male, 24 years, intervention arm, Mulago)
- Overcoming the stigma and discrimination (how to disclose the HIV status)
  - "Teach us how to manage stigma. I have seen some people...who insult the positive ones. They keep saying so and so is sick and many more words...teach us how to respond to other people in case they stigmatize us." (IDI, female, 19 years, intervention arm, Mulago)
- Sexual maturation (how to manage body odor)
  - "Young people these days have bad odours, so call for life can come in and sensitize on how to ensure good body hygiene. Teach us how to manage and prevent that foul smell." (IDI, male, 24 years, intervention, Mulago)
- Explanation on viral load
  - "Personally, I take a while to know about my viral load. They can tell us whether it is high or low and its implications." (IDI, male, 24 years, intervention arm, Mulago)
- Dangers of alcohol use and drug abuse
  - "Advise youth not to take alcohol and drugs. There are some who use cigarettes and marijuana, there is need to teach them avoidance of harmful habits like alcohol especially when one is on drugs" (IDI, female, 24 years, intervention arm, Mulago)
- Drug side effects
  - "Teach us when and how to swallow the drugs. If there is any side effect, explain how to manage them. I think that education can be helpful." (IDI, male, 24 years, intervention arm, Mulago)
- Basic antiretroviral therapy (ART) education and emphasis on ART adherence
  - "Re-emphasize ART adherence...if you don't swallow your pills you will get such and such problems or if you don't swallow your drugs on time, you will get such and such problem." (IDI, female, 23 years, intervention arm, Mulago)
- How to overcome low self-esteem, self-pity, and stress management
  - "They should teach me how I should conduct myself in public. I have that problem of low self-esteem." (IDI, female, 19 years, intervention arm, Mulago)
- Income generation activities or value of vocational skills (art and crafts)
  - "Teach us value of income generation activities, teach us vocational skills such as Art and crafts, and Candle making to earn a living." (FGD, female, 18-24 years, intervention, Kasangati)
- Pregnancy and breastfeeding options
  - "Pregnant women need antenatal reminders through call for life system. Just the way routine appointment reminders are done by system." (IDI, female, 22 years, intervention arm, Kasangati)
- Family planning
  - "They should teach us about family planning...how it works some of us who want to access family planning. We were told they have to first check our blood to determine the appropriate method...personally I was on FP but I got an unintended pregnancy." (IDI, female, 22 years, intervention arm, Mulago)
- Herbal treatment (risks and benefits)
  -

"Add tips like, you don't have to use herbal remedies educates me about the risks and benefits of herbs that I was ignorant about." (IDI, female, 23 years, intervention arm, Mulago)

- Good nutrition
  - "You can decide that today I want good nutrition tips and they tell you the food categories that will improve the CD4." (IDI, male, 19 years, intervention, Kasangati)

Daily Pill Reminder Call

Youth, mainly from Kasangati, found that the twice daily remainder call that came through at an agreed specific time was found to be useful in promoting adherence:

> *What I love the most is the pill reminders because before I joined, I used to take when it's already past the time, but now I take my drugs on time. That's what I like most about it.* [IDI, female, 22 years, intervention arm, Kasangati]

Barriers to and Challenges of the CFLU System

Technical Issues With the System

IDIs and FGD elicited challenges related to the tool. According to most young adults, these were mainly the temporary technical problems that they faced during an upgrade to a newer version. This led to interruptions in calls and the system calling beyond the time scheduled.

Youth from both sites, mainly those from Infectious Diseases Institute Clinic, reported blocked pin codes, the failure of the system to complete outbound health tips, and poor pin recognition (nonresponse after entering a pin code). Most youth reported that the irregularities or inconsistencies of the system were frustrating:

> *...it was recently messed up a bit. I had stopped picking up the phone calls because every time the call comes through and I put in the pin number, it goes off.* [FGD, female,18-24 years, intervention arm, Mulago]

Poor Access to Mobile Phones

Lost phone problems and battery issues made it difficult for some young people to access CFLU:

> *There was a time when my phone was stolen which required me to buy another replacement. I was off air for one month. I found it a problem because the system was helpful...* [IDI, male, 19 years, intervention arm, Mulago]

Fear of the HIV-Related Stigma

There was a shared perspective from IDIs and FGD young adults from both sites about the fear of stigma and discrimination associated with HIV. The perception that HIV is a fatal disease associated with promiscuity led to low self-esteem and poor adherence among youth:

> *Us youths...are shy...proud and boastful with a lot of pride. We don't want others to know we are positive. If they receive call and accidentally insert the pin code...if accidentally the phone is in loud it can be shameful, one can stop medication...that leads to opportunistic infections.* [IDI, male, 19 years, intervention, Kasangati]

> *Youth fear to be stigmatized. For instance, I had a friend but once she found out about my status, she stopped sharing food with me on the same plate, drinking from the same cup, because she did not want to catch HIV.* [IDI, 24 years, female, intervention, Mulago]

Majority of young adults from the interviews feared being seen taking medicine or answering or responding to a CFLU call because it would be associated with HIV. Few youth feared to be found by their peers or neighbors during their routine clinic visits. Moreover, a few young adults complained of noisy pill tins because they attracted attention. They also worried about their friends spreading rumors about their HIV status:

> *For me I had a bad chance, the people I found were very ill, I asked for Wednesday, that's the day, I found that lady who spread rumours. I felt small I was not comfortable even at my work place it still affects me.* [FGD, female, 18-24 years, intervention, Kasangati]

> *People pin point at us; the youth, so we fear to take pills. The people may be in the surroundings be it at work or at home, someone may come and rub it your face that you are HIV+. It has happened to me before and I stopped taking my pills.* [IDI, 24 years, female, intervention arm, Mulago]

Suggestions and Recommendations From Youth to Improve CFLU

Promote Peer Support Meeting

Although there was a positive response to the tool, most young people provided suggestions to improve CFLU. They requested extra support to scale up the CFLU system, such as promotion of peer support meetings among youth with HIV infection on CFLU. This would help them break stigma trends by sharing experiences and attaining peer mentorship. Majority of youth said that peer support meetings would promote youth's economic capacity to become financially independent. This could be led by empowering them to start an income generation activity as a way of addressing some of the adherence challenges:

> *Teach us vocational skills like arts and crafts, through peer support meetings we could be given vocational skills of which we can learn to make products for sale to earn a living...CFLU could help and train us who can learn different skills...for transport to hospital*

and self-care. [IDI, 22 years, intervention arm, Kasangati]

Community Sensitization About HIV and CFLU

Young adults, mainly from Kasangati, suggested community sensitization about HIV to create awareness. They said extensive community sensitization about HIV reduced the stigma associated with HIV, which could subsequently make it easier for youth to adhere perfectly to their drugs:

...get a day and invite all people whether positive or not to retest for HIV then enrol them on call for life. [IDI, female, 22 years, intervention arm, Kasangati]

Establish Youth-Friendly HIV Services and Scale Up CFLU

Youth requested for additional health tips, setting up of a designated space and staff for CFLU youth, resolving technical issues, and supporting youth to overcome stigma and discrimination, as summarized in Textbox 2 and Table 3.

Textbox 2. Suggestions by young adults to improve the use of the Call for Life Uganda (CFLU) system.

Suggestions and illustrative quotes

- Encourage young people to maintain the same phone line
  - "Many young adults change their phone lines and fail to inform the health providers about it. Tell them stick to one line or in case they change, they should inform the health providers about the new phone lines so that they don't miss the call." (Focus group discussion [FGD], female, 18-24 years, intervention arm, Kasangati)
- Resolve technical errors
  - "In order to maintain the standard, ensure the call for life staff resolve the technical issues about the system. Staff should ensure the daily calls come through and call at the agreed time. For instance, I take my pills at 10pm but they call me 30 minutes later daily." (In-depth interview [IDI], female, 24 years, intervention arm, Mulago)
- Make CFLU appealing to young people and explain the benefits
  - "Us young people need a TV and games at the clinic waiting areas to occupy us when we arrive as we wait for the call for life doctors to welcome us, health educate us and treat us." (FGD, female, young adults, 18-24 years, intervention arm, Kasangati)
- Scale up CFLU to all adults and youth in other facilities
  - "First, enrol all the adults and young people on call for life system so that they can benefit from it...when they use the system it will help remind them to take their daily medicine on time." (IDI, male, 19 years, intervention arm, Kasangati)
- Create champions of CFLU to influence other youth to adopt CFLU
  - "Organise a day for young people on call for life...to share experiences with fellow youth who are not experienced on call for life system, us on the system are the champions. We can learn from each other about adherence, avoidance of self-pity and benefits of ART...Another thing it is encouraging for us young people when we associate with fellow peers, we can influence youths to join the system. (IDI, female, 19 years, intervention arm, Kasangati)
- Establish a designated space and day for CFLU
  - "We need a designated space where we can wait. It is a big challenge for us when we come here. We do not have a call for life premise and usually we met the wrong doctors when we come here. You hear young people say, there is a doctor who called me; I don't know how to locate the doctor." (FGD, female, 18-24 years, intervention arm, Kasangati)
- Establish a designated staff to handle young adults enrolled on CFLU at Kasangati Health facility.
  - "The call for life staff people should let us know the focal person to talk to when we come here...It's by luck to find the right doctor and room to go to...we should get a permanent call for life space, so that we know that once we come we go right away to that place or room, Instead of wasting time looking around." (FGD, female, 18-24 years, intervention arm, Kasangati)
- Sensitize all patients about the CFLU system and do it as a team—doctors, nurses, and patients. Make announcements on television and radio
  - "You have to work as a team, doctors, nurses, counselors, tell the patients about the benefits of the system. If you sensitise more people on TV or radio, very many people will come up and join the system. It's just about creating awareness regarding how the system works and its related benefits." (IDI, 24 years, intervention arm, Mulago)
- Support young people to overcome the stigma
  - "Teach us (young people) how to overcome stigma and respond to people who point fingers at us." (FGD, female, 18-24 years, intervention arm, Kasangati)
- Combine research study appointments with routine general HIV clinic so that patients do not come twice
  - "Suggest that doctors Combine study appointments with general clinic so that patients do not come twice. The return appointment for call for life system should be recorded in the system and in the general file...When one comes for call for life clinic visit...prior to return appointment, they should supply drugs once for all to avoid confusion." (FGD, female, 18-24 years, intervention arm, Kasangati)

Table 3. Suggestions and adherence mitigation strategies from young adults through the Call for Life Uganda (CFLU) system.

Subject and factors hindering ART[a] adherence	Strategies proposed through CFLU
Personal barrier	
Forgetfulness secondary to:	Mitigation
Busy work schedule	Daily pill reminder function
School schedules	Daily pill reminder function
Sports activities	Daily pill reminder function
Dancing and alcohol	Daily pill reminder function
Psychosocial	
Nondisclosure of serostatus	Health tips call; psychosocial support
Stigma (fear to be seen taking pills and raising suspicion from people)	Health tips on Psychosocial support
Mental issues (low self-esteem as a result of having HIV)	Health tips on Psychosocial support
Pride or irresponsibility	Health tips on Psychosocial support
Misconception about HIV results (go off drugs when they improve and the viral load is undetected)	Counseling
Lack of psychosocial support	Treatment buddies and psychosocial support
Pill related	
Drug-related side effects, for example, dizziness and big size of the pill	Symptom reporting and management through the CFLU system
Sound or noise by pill hitting the pill bottles	Repacking of pills
Socioeconomic status	
Missed clinic appointments	Clinic visit call reminders
Missed clinic appointments because of lack of transport	Joining community drug distribution points

[a]ART: antiretroviral therapy.

Discussion

Principal Findings

In this study, we sought to describe the acceptability of a mobile phone support tool to promote adherence to ART in young adults in an RCT. Understanding the acceptability of a mobile phone support tool to promote adherence to ART in young adults was important to enhance adherence intervention strategies. Four themes emerged from interviews with young adults, including the positive attitude toward CFLU, which included improvement of medication adherence, management of problems of forgetfulness, management of stigma-related issues, and psychosocial support. Negative attitudes toward CFLU were reported under barriers to and challenges of the CFLU theme, including technical issues, poor access to mobile phones, fear of stigma, and financial constraints. One theme was about the components of CFLU and suggestions and recommendations from youth to improve CFLU. This study reveals that the CFLU system supports adherence, which is a critical challenge for youth on ART. The most common barriers to ART adherence mentioned by the young adults in this study included forgetfulness, nondisclosure of serostatus, ART-related side effects, stigma, and pill burden. Forgetfulness is a major factor previously pointed out in other mHealth studies as a barrier to ART adherence [23]. Majority of young adults in this study said that CFLU resolved most of the adherence challenges such as forgetfulness, missed pills and clinic visits, and fear of stigma.

Fear of stigma and discrimination associated with HIV was reported as a challenge to CFLU, which is similar to previous mHealth studies [24]. Health tips emphasized the importance of disclosure to avoid stigma that would affect drug adherence. Patients suggested the desire for CFLU to continue educating youth about HIV and how to overcome stigma through continuous health tips.

Results revealed that CFLU resolved most of the barriers to adherence, such as forgetfulness. A major barrier to ART adherence was also tied closely with daily routine, as reported in some other study [25]. There is a need to make CFLU appealing to the youth so that it can be scaled up to eradicate barriers to ART initiation and adherence, such as fear of disclosing HIV status to partners, drug-related factors (side effects and the big size of the tablet), and HIV stigma [26].

This is in agreement with mHealth interventions that were first deployed in noncommunicable diseases and later used in infectious diseases [27]. Evidence reviews suggest that mHealth interventions delivered in low-income and middle-income countries can be effective in improving health outcomes for people living with chronic diseases [28]. A systematic review of mHealth interventions for monitoring chronic disease by Watkins et al [29] found articles on the monitoring of

hypertension, stroke, and people living with HIV from Kenya, Pakistan, Honduras, Mexico, and South Africa. The 6 components of mHealth found in all 4 interventions included reminders, patient observation of health state, motivational education or advice, provision of support communication, targeted actions, and praise and encouragement [29]. Communicating with young adults about health issues and adherence and explaining the possible side effects are important in enhancing their informed decision making.

Concerns about CFLU reported ranged from technical issues, poor access to mobile phones, fear of stigma, and financial constraints. The disadvantages of CFLU reported by young adults were similar to those reported in other studies that reported the cost and convenience of SMS text message, given that its low cost is well suited for supporting the treatment of conditions managed over extended periods compared with interactive voice calls [23]. In the WelTel Kenya 1 trial, SMS intervention was considered inexpensive, and each SMS costed approximately US $0.05, equivalent to US $20 per 100 patients per month, and follow-up voice calls averaged US $3.75 per nurse per month [19]; however, SMS can only be applicable in a population with some literacy [30]. One study in Zanzibar showed that behavioral change was significantly higher with pushed SMS enrollees than with voice messaging enrollees [31]; this was not assessed in our study. However, of the 300 young adults enrolled in the CFLU mHealth intervention arm, only 2 opted for SMS text messaging and were not among the young adults.

Lost phone problems and battery issues made it difficult for some young people to access CFLU, and few young adults preferred messages that they could retrieve after restoring the phone battery. Similar studies report that message delivery rates are far more successful among SMSs than among voice enrollees. Pushing voice messages to clients with personal phones is a complex process that requires the client to answer the phone at the time of delivery, whereas SMS messages can be delivered at any time, including when the phone is turned off. This is the major reason that SMS messages pushed to personal phones have a higher delivery success rate than voice messages [32]. The disadvantage of SMS text messaging reported by other studies includes breaches of confidentiality, and the majority of the patients have a fear of revealing their HIV serostatus, which makes voice calls superior to SMSs unless messages are coded.

The use of the individualized secret pin code enhanced the confidentiality and privacy of most young adults on CFLU. However, blocked pin codes and poor pin recognition of the system hindered the use of the system for most young adults. In a similar study in western Uganda, pin-protected messages reduced the odds of clinic return, as the use of pins was a challenge; hence, they often missed the message alerting them for clinic appointments [30]. Our adolescents advised empowering youth to overcome stigma and setting up a customer help desk at the health facility to resolve technical issues related to CFLU. Other studies suggest the involvement of end users during the development of mHealth apps.

Limitations

At the time of these interviews (September 2017), the CFLU system was experiencing technical issues following a newer software release; therefore, the calls kept dropping for both outbound and inbound calls. This was resolved first by contacting the software developer and consultant, halting calls, and temporarily turning off the system until the causes for technical issues were investigated and resolved. Automated alerts were developed to notify the principal investigator, the study coordinator, and the information system team on system errors.

During the same period, there was a national deadline by the Uganda Communication Commission to register all phone SIM cards, and some phones were cut off during this period. Young adults bought newly registered phone lines for CFLU.

Conclusions

The CFLU system can support adherence, despite some of the temporary technical issues. Enhancing adherence to ART using the CFLU system addresses the challenges reported by young people. The CFLU system is user friendly, acceptable, and a feasible strategy to monitor and improve adherence of patients in resource-limited settings.

Acknowledgments

The authors are grateful to the participants and staff at the Infectious Diseases Institute and Kasangati HIV clinics for their participation in this study. This study and the Academy for Health Innovations, Uganda, were funded by Janssen, the Pharmaceutical Companies of Johnson & Johnson, as part of its commitment to global public health through collaboration with the Johnson & Johnson Corporate Citizenship Trust. The Straight Talk Foundation, Uganda, supported the development, recording, and piloting of the HIV health tip content used in this study.

Authors' Contributions

AT led the data collection and data analysis and drafted the first manuscript. RPR led the study design and implementation and contributed to manuscript writing and reviewing. ABN and EO led the study implementation and data collection and contributed to manuscript writing and reviewing. RK, AK, BC, JW, MSN, ETK, ML, and RK contributed to the study design and manuscript reviewing until the final version approval.

Conflicts of Interest

RPR discloses that Infectious Diseases Institute receives research funding from Janssen, the Pharmaceutical Companies of Johnson and Johnson for work on Call for Life and other research projects.

Multimedia Appendix 1
The call flow from the system to the end user, with functionalities for daily pill reminder calls, clinic visit reminders, health tips, and the interactive guide and options until a call ends.
[DOCX File , 239 KB - mhealth_v9i6e17418_app1.docx]

Multimedia Appendix 2
Flowchart of young adults involved in the qualitative study.
[DOCX File , 21 KB - mhealth_v9i6e17418_app2.docx]

References

1. UNAIDS data 2019. UN Joint Programme on HIV/AIDS (UNAIDS). 2019. URL: https://www.unaids.org/en/resources/documents/2019/2019-UNAIDS-data [accessed 2019-12-15]
2. 2018 Statistical Abstract. Uganda Bureau of Statistics. 2018. URL: https://www.ubos.org/wp-content/uploads/publications/05_2019STATISTICAL_ABSTRACT_2018.pdf [accessed 2019-12-10]
3. Uganda Demographic and Health Survey. Uganda Bureau of statistics. 2016. URL: https://dhsprogram.com/pubs/pdf/FR333/FR333.pdf [accessed 2019-12-02]
4. Uganda population-based HIV impact assessment (UPHIA) 2016-2017. Ministry of Health, Uganda. 2019 Jul. URL: https://phia.icap.columbia.edu/wp-content/uploads/2019/07/UPHIA_Final_Report_Revise_07.11.2019_Final_for-web.pdf [accessed 2019-12-16]
5. Nachega JB, Hislop M, Dowdy DW, Chaisson RE, Regensberg L, Maartens G. Adherence to nonnucleoside reverse transcriptase inhibitor-based HIV therapy and virologic outcomes. Ann Intern Med 2007 Apr 17;146(8):564-573. [doi: 10.7326/0003-4819-146-8-200704170-00007] [Medline: 17438315]
6. Sethi AK, Celentano DD, Gange SJ, Moore RD, Gallant JE. Association between adherence to antiretroviral therapy and human immunodeficiency virus drug resistance. Clin Infect Dis 2003 Oct 15;37(8):1112-1118. [doi: 10.1086/378301] [Medline: 14523777]
7. Harrigan P, Hogg R, Dong W, Yip B, Wynhoven B, Woodward J, et al. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. J Infect Dis 2005 Feb 01;191(3):339-347. [doi: 10.1086/427192] [Medline: 15633092]
8. Nachega J, Hislop M, Dowdy DW, Lo M, Omer SB, Regensberg L, et al. Adherence to highly active antiretroviral therapy assessed by pharmacy claims predicts survival in HIV-infected South African adults. J Acquir Immune Defic Syndr 2006 Sep;43(1):78-84. [doi: 10.1097/01.qai.0000225015.43266.46] [Medline: 16878045]
9. Whiteley L, Brown L, Lally M, Heck N, van den Berg JJ. A mobile gaming intervention to increase adherence to antiretroviral treatment for youth living with HIV: development guided by the information, motivation, and behavioral skills model. JMIR Mhealth Uhealth 2018 Apr 23;6(4):e96 [FREE Full text] [doi: 10.2196/mhealth.8155] [Medline: 29685863]
10. Naomi W. Determinants of adherence to antiretroviral among HIV positive adolescents at comprehensive care clinic, Gertrude's Children Hospital, Nairobi, Kenya. Am J Nurs Sci 2018;7(1):23. [doi: 10.11648/j.ajns.20180701.13]
11. Becker SL, Dezii CM, Burtcel B, Kawabata H, Hodder S. Young HIV-infected adults are at greater risk for medication nonadherence. MedGenMed 2002 Jul 30;4(3):21 [FREE Full text] [Medline: 12466764]
12. Reda AA, Biadgilign S. Determinants of adherence to antiretroviral therapy among HIV-infected patients in Africa. AIDS Res Treat 2012;2012:574656 [FREE Full text] [doi: 10.1155/2012/574656] [Medline: 22461980]
13. The Uganda National Household Survey 2016/2017. Uganda Bureau of Statistics (UBOS). 2018. URL: https://www.ubos.org/wp-content/uploads/publications/03_20182016_UNHS_FINAL_REPORT.pdf [accessed 2019-12-06]
14. Mitchell KJ, Bull S, Kiwanuka J, Ybarra ML. Cell phone usage among adolescents in Uganda: acceptability for relaying health information. Health Educ Res 2011 Oct;26(5):770-781 [FREE Full text] [doi: 10.1093/her/cyr022] [Medline: 21536715]
15. Kunutsor S, Walley J, Katabira E, Muchuro S, Balidawa H, Namagala E, et al. Using mobile phones to improve clinic attendance amongst an antiretroviral treatment cohort in rural Uganda: a cross-sectional and prospective study. AIDS Behav 2010 Dec;14(6):1347-1352. [doi: 10.1007/s10461-010-9780-2] [Medline: 20700644]
16. 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] [Medline: 25881059]
17. Ammassari A, Trotta MP, Shalev N, Tettoni MC, Maschi S, Di Sora F, et al. Timed short messaging service improves adherence and virological outcomes in HIV-1-infected patients with suboptimal adherence to antiretroviral therapy. J Acquir Immune Defic Syndr 2011 Dec 01;58(4):e113-e115. [doi: 10.1097/QAI.0b013e3182359d2a] [Medline: 22033235]

18. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. BMC Health Serv Res 2015 Jan 22;15:33 [FREE Full text] [doi: 10.1186/s12913-014-0654-6] [Medline: 25609559]
19. Lester R, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
20. Haberer JE, Musiimenta A, Atukunda EC, Musinguzi N, Wyatt MA, Ware NC, et al. Short message service (SMS) reminders and real-time adherence monitoring improve antiretroviral therapy adherence in rural Uganda. AIDS 2016 May 15;30(8):1295-1300 [FREE Full text] [doi: 10.1097/QAD.0000000000001021] [Medline: 26760452]
21. Sidney K, Antony J, Rodrigues R, Arumugam K, Krishnamurthy S, D'souza G, et al. Supporting patient adherence to antiretrovirals using mobile phone reminders: patient responses from South India. AIDS Care 2012;24(5):612-617. [doi: 10.1080/09540121.2011.630357] [Medline: 22150088]
22. Mack N, Woodsong C, MacQueen KM, Guest G, Namey E. Focus groups. In: Qualitative Research Methods: A Data Collector's Field Guide. Research Triangle Park, North Carolina 27709, USA: Family Health International; 2005:51-82.
23. Garofalo R, Kuhns LM, Hotton A, Johnson A, Muldoon A, Rice D. A randomized controlled trial of personalized text message reminders to promote medication adherence among HIV-positive adolescents and young adults. AIDS Behav 2016 May;20(5):1049-1059 [FREE Full text] [doi: 10.1007/s10461-015-1192-x] [Medline: 26362167]
24. MacDonell K, Naar-King S, Huszti H, Belzer M. Barriers to medication adherence in behaviorally and perinatally infected youth living with HIV. AIDS Behav 2013 Jan;17(1):86-93 [FREE Full text] [doi: 10.1007/s10461-012-0364-1] [Medline: 23142855]
25. Murphy DA, Sarr M, Durako SJ, Moscicki AB, Wilson CM, Muenz LR, Adolescent Medicine HIV/AIDS Research Network. Barriers to HAART adherence among human immunodeficiency virus-infected adolescents. Arch Pediatr Adolesc Med 2003 Mar;157(3):249-255. [doi: 10.1001/archpedi.157.3.249] [Medline: 12622674]
26. Buregyeya E, Naigino R, Mukose A, Makumbi F, Esiru G, Arinaitwe J, et al. Facilitators and barriers to uptake and adherence to lifelong antiretroviral therapy among HIV infected pregnant women in Uganda: a qualitative study. BMC Pregnancy Childbirth 2017 Mar 21;17(1):94 [FREE Full text] [doi: 10.1186/s12884-017-1276-x] [Medline: 28320347]
27. Ali EE, Chew L, Yap KY. Evolution and current status of mhealth research: a systematic review. BMJ Innov 2016 Jan 05;2(1):33-40. [doi: 10.1136/bmjinnov-2015-000096]
28. Watkins JA, Goudge J, Gómez-Olivé FX, Huxley C, Dodd K, Griffiths F. mHealth text and voice communication for monitoring people with chronic diseases in low-resource settings: a realist review. BMJ Glob Health 2018;3(2):e000543 [FREE Full text] [doi: 10.1136/bmjgh-2017-000543] [Medline: 29527356]
29. Watkins JO, Goudge J, Gómez-Olivé FX, Griffiths F. Mobile phone use among patients and health workers to enhance primary healthcare: a qualitative study in rural South Africa. Soc Sci Med 2018 Feb;198:139-147 [FREE Full text] [doi: 10.1016/j.socscimed.2018.01.011] [Medline: 29335160]
30. Siedner MJ, Santorino D, Haberer JE, Bangsberg DR. Know your audience: predictors of success for a patient-centered texting app to augment linkage to HIV care in rural Uganda. J Med Internet Res 2015 Mar 24;17(3):e78 [FREE Full text] [doi: 10.2196/jmir.3859] [Medline: 25831269]
31. Lund S, Hemed M, Nielsen BB, Said A, Said K, Makungu MH, et al. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomised controlled trial. Br J Obstet Gynaecol 2012 Sep;119(10):1256-1264. [doi: 10.1111/j.1471-0528.2012.03413.x] [Medline: 22805598]
32. Crawford J, Larsen-Cooper E, Jezman Z, Cunningham SC, Bancroft E. SMS versus voice messaging to deliver MNCH communication in rural Malawi: assessment of delivery success and user experience. Glob Health Sci Pract 2014 Feb;2(1):35-46 [FREE Full text] [doi: 10.9745/GHSP-D-13-00155] [Medline: 25276561]

Abbreviations

ART: antiretroviral therapy
CfL: connect for life
CFLU: Call for Life Uganda
FGD: focus group discussion
IDI: in-depth interview
IVR: interactive voice response
mHealth: mobile health
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 13.12.19; peer-reviewed by S Harrison, B Oh, J Knox, J Truong; comments to author 26.08.20; revised version received 19.10.20; accepted 07.05.21; published 14.06.21.

Please cite as:
Twimukye A, Bwanika Naggirinya A, Parkes-Ratanshi R, Kasirye R, Kiragga A, Castelnuovo B, Wasswa J, Nabaggala MS, Katabira E, Lamorde M, King RL
Acceptability of a Mobile Phone Support Tool (Call for Life Uganda) for Promoting Adherence to Antiretroviral Therapy Among Young Adults in a Randomized Controlled Trial: Exploratory Qualitative Study
JMIR Mhealth Uhealth 2021;9(6):e17418
URL: https://mhealth.jmir.org/2021/6/e17418
doi:10.2196/17418
PMID:34121665

Original Paper

Telerehabilitation for Lung Transplant Candidates and Recipients During the COVID-19 Pandemic: Program Evaluation

Lisa Wickerson[1,2], BScPT, MSc, PhD; Denise Helm[1,2], BScPT; Chaya Gottesman[1,2], BScPT; Dmitry Rozenberg[1,3], MD, PhD; Lianne G Singer[1,3], MD; Shaf Keshavjee[1,4], MD, MSc; Aman Sidhu[1,3], MD, MSc

[1]Toronto Lung Transplant Program, University Health Network, Toronto, ON, Canada

[2]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada

[3]Department of Medicine, University of Toronto, Toronto, ON, Canada

[4]Department of Surgery, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Lisa Wickerson, BScPT, MSc, PhD
Toronto Lung Transplant Program
University Health Network
585 University Ave
Toronto, ON, M5G 2N2
Canada
Phone: 1 437 229 0049
Fax: 1 416 340 4772
Email: lisa.wickerson@uhn.ca

Abstract

Background: The COVID-19 pandemic resulted in a rapid shift from center-based rehabilitation to telerehabilitation for chronic respiratory disease and lung transplantation due to infection control precautions. Clinical experience with this delivery model on a large scale has not been described.

Objective: The aim of this study is to describe usage and satisfaction of providers and lung transplant (LTx) candidates and recipients and functional outcomes following the broad implementation of telerehabilitation with remote patient monitoring during the first wave of the COVID-19 pandemic.

Methods: This study was a program evaluation of providers, LTx candidates, and early LTx recipients who used a web-based, remote monitoring app for at least four weeks between March 16 and September 1, 2020, to participate in telerehabilitation. Within-subjects analysis was performed for physical activity, Self-efficacy For Exercise (SEE) scale score, aerobic and resistance exercise volumes, 6-minute walk test results, and Short Physical Performance Battery (SPPB) results.

Results: In total, 78 LTx candidates and 33 recipients were included (57 [51%] males, mean age 58 [SD 12] years, 58 [52%] with interstitial lung disease, 34 [31%] with chronic obstructive pulmonary disease). A total of 50 (64%) LTx candidates and 17 (51%) LTx recipients entered ≥10 prescribed exercise sessions into the app during the study time frame. In addition, 35/42 (83%) candidates agreed the app helped prepare them for surgery and 18/21 (85%) recipients found the app helpful in their self-recovery. The strongest barrier perceived by physiotherapists delivering the telerehabilitation was patient access to home exercise and monitoring equipment. Between the time of app registration and ≥4 weeks on the waiting list, 26 LTx candidates used a treadmill, with sessions increasing in mean duration (from 16 to 22 minutes, *P*=.002) but not speed (from 1.7 to 1.75 mph, *P*=.31). Quadriceps weight (pounds) for leg extension did not change (median 3.5, IQR 2.4-5 versus median 4.3, IQR 3-5; *P*=.08; n=37). On the Rapid Assessment of Physical Activity questionnaire (RAPA), 57% of LTx candidates scored as active, which improved to 87% (*P*=.02; n=23). There was a decrease in pretransplant 6-minute walk distance (6MWD) from 346 (SD 84) meters to 307 (SD 85) meters (*P*=.002; n=45) and no change in the SPPB result (12 [IQR 9.5-12] versus 12 [IQR 10-12]; *P*=.90; n=42). A total of 9 LTx recipients used a treadmill that increased in speed (from 1.9 to 2.7 mph; *P*=.003) between hospital discharge and three months posttransplant. Quadriceps weight increased (3 [IQR 0-3] pounds versus 5 [IQR 3.8-6.5] pounds; *P*<.001; n=15). At three months posttransplant, 76% of LTx recipients scored as active (n=17), with a high total SEE score of 74 (SD 11; n=12). In addition, three months posttransplant, 6MWD was 62% (SD 18%) predicted (n=8).

Conclusions: We were able to provide telerehabilitation despite challenges around exercise equipment. This early experience will inform the development of a robust and equitable telerehabilitation model beyond the COVID-19 pandemic.

(JMIR Mhealth Uhealth 2021;9(6):e28708) doi:10.2196/28708

KEYWORDS

telerehabilitation; lung; transplant; rehabilitation; COVID-19; usage; satisfaction; app; outcome; telemedicine

Introduction

Lung transplant (LTx) candidates exhibit reduced aerobic exercise capacity, low physical activity levels, and muscle weakness, which diminishes further in the early posttransplant period [1-3]. Reduced aerobic capacity is a strong predictor of mortality pretransplant and is associated with worse posttransplant health outcomes, including longer length of hospital stay and decreased survival [4-6]. Conversely, greater levels of physical activity, muscle strength, and exercise capacity after LTx are associated with reduced development of cardiovascular comorbidities and better longer-term health outcomes such as quality of life [1,2,7]. Pretransplant exercise training is therefore recommended to optimize the benefits of transplantation [8].

Our center's mandatory pretransplant and posttransplant rehabilitation has historically been center-based, requiring patients to travel or relocate to participate. This added to treatment burden, and our LTx program was exploring ways to support patients closer to home. The COVID-19 pandemic resulted in a rapid adoption of virtual care including telerehabilitation [9-11]. Telerehabilitation is defined as the delivery of rehabilitation at a distance using a variety of information communication technologies. Models include real-time videoconferencing with telemonitoring of individuals or groups, and asynchronous web-, app-, or phone-based models with remote monitoring of biometrics [12].

To date, little is known about the feasibility, efficacy, and effectiveness of telerehabilitation in LTx candidates and recipients. A web-based platform delivering 8 weeks of telerehabilitation early following hospitalization was safe and associated with increased functional exercise capacity, balance, lower limb strength, and physical activity in 4 LTx recipients [13]. The same research group reported that an 8-week home rehabilitation program in one LTx candidate improved functional outcomes [14]. A pilot study of home rehabilitation to decrease physical frailty in 13 LTx candidates using a mobile app in addition to weekly phone check-ins was found to be safe and feasible [15]. These studies included small groups of patients, and consequently there is a lack of clinical experience with this delivery model on a large scale and especially during a pandemic. Clinical trials comparing telerehabilitation and center-based rehabilitation have included an initial in-person exercise assessment to determine a safe and effective exercise prescription; however, this has not always been feasible during the COVID-19 pandemic [16].

In 2019, our LTx program planned a 2-year, interdisciplinary clinical project to procure and trial a commercially available, customizable, web-based, remote care platform (the Vivify Health app) to support patients in preparing for and recovering from LTx closer to home through monitoring, telerehabilitation, and communication by the surgical, medical, and rehabilitation teams. Features of this platform include an online patient education library, prompts and reminders, satisfaction and symptom surveys, biometric data monitoring, alerts triggered on the clinician's dashboard view if entered biometrics are outside of set parameters or if a question is answered with a clinically relevant response, personalized care plans including a daily individualized exercise pathway that is filled out at the time of exercise, and asynchronous in-app texting and embedded secure videoconferencing between patients and the health care team. The project rolled out in January 2020 targeting 10 patients; however, in response to the pandemic in mid-March 2020, all LTx candidates and LTx recipients less than three months posttransplant were approached to register for the app to enable mobile asynchronous communication with the health care team, virtual visits with clinicians, telerehabilitation, and remote monitoring to limit on-site hospital visits.

The aim of this program evaluation was to examine patient and provider satisfaction, usage of the exercise pathway, and exercise and physical functional outcomes of a large group of LTx candidates and recipients who used the app during the first wave of the COVID-19 pandemic in order to inform ongoing improvements to a telerehabilitation model.

Methods

A program evaluation of LTx candidates and recipients who used the remote monitoring app platform for telerehabilitation for at least four weeks between March 16 and September 1, 2020, was completed. In this study, four weeks was used as a cutoff to allow time for adjustments to exertional oxygen prescription and to observe the anticipated benefits of rehabilitation. Usage was tracked in the app (number of exercise sessions patients entered, number of times educational resources were accessed, and number of video visits performed by physiotherapists). Satisfaction with the app was measured through a survey that was sent to patients two weeks after registration pre-LTx and three months post-LTx and was not specific to rehabilitation (Multimedia Appendix 1). A survey was also sent in September 2020 to physiotherapists delivering telerehabilitation (Multimedia Appendix 2). Home exercise and monitoring equipment access was collected in a survey administered in the app at the time of app registration. Exertional oxygen use was tracked in the app through patient self-report. Physical activity and self-efficacy for exercise were measured using the Rapid Assessment of Physical Activity (RAPA) questionnaire [17] and Self-Efficacy for Exercise (SEE) scale [18], which were sent to LTx candidates in the app at baseline after app registration and four weeks later, and to LTx recipients three months posttransplant. Exercise volumes were tracked in the app and monitored by physiotherapists. Exercise data for LTx candidates were taken at baseline (first week after app registration) and repeated at the last rehabilitation entry in the app during the study time frame. Exercise data for LTx recipients were taken at baseline (one week after hospital

discharge) and at three months posttransplant. Between March 16 and June 30, 2020, there were very few on-site visits for functional and exertional oxygen assessment due to pandemic restrictions, and all exercise interventions were performed remotely at home. Between July 1 and August 31, 2020, there was an increase in on-site functional and oxygen assessments, and 1-3 initial sessions for exercise instruction, but the majority of rehabilitation was app-guided unsupervised home-based exercise with patient self-monitoring and manual entry into the app. Most external pulmonary rehabilitation programs and communal gyms were closed. An aerobic and resistance exercise program was individually tailored to the patients' exertional oxygen requirements, disease stability, functional capacity, and access to home exercise equipment. The exercises were prescribed at least three times a week for the duration of the wait time pretransplant and between hospital discharge and three months posttransplant. Telehealth support was provided by the physiotherapist by phone, video, asynchronous texting, and remote monitoring. Declines in exercise capacity, progression of symptoms, and increased exertional oxygen requirements were regularly discussed with the transplant medical team. Standard functional outcomes included the 6-minute walk test (6MWT) and the Short Physical Performance Battery (SPPB), which were performed at the start of rehabilitation after listing for lung transplant and every three months on the waiting list pretransplant. The 6MWT was also repeated three months posttransplant when on-site visits were permitted. Due to the urgency of transitioning all patients to telerehabilitation during the pandemic to avoid on-site visits for patient and staff safety—and a consequent lack of a control group undergoing traditional in-person rehabilitation during the same time period—we report functional data of our center-based rehabilitation program from research studies conducted between 2010 and 2019.

Exclusion criteria for app registration included no access to a supported model of smartphone or tablet, no phone data alongside unreliable or limited Wi-Fi, and patients unable or unwilling to use the technology, although patients could choose to have a proxy caregiver register and access the app for them. Patients did not use the app for rehabilitation while they were admitted to hospital pretransplant or posttransplant.

The app is a browser-based solution provided through a third-party vendor that is licensed by Health Canada, has a Class I Medical Device Establishment License, and stores all data on remote servers in Canada. Safety measures included assessment from our institutional privacy and security departments, regular penetration testing and data encryption, clinician access though security assertion markup language integration, and patient access through two-factor authentication. All use within the app is auditable and time-stamped. Patients provided written consent on an end-user license agreement to allow the data they entered into the app to be used for clinical care and quality improvement. This program evaluation was reviewed and approved by our institution's Quality Improvement Review Committee.

For statistical analysis, normality of the data was checked using the Shapiro-Wilk test. Continuous variables were summarized as mean (SD) or median (IQR). Categorical variables were summarized as counts and percentages. Paired *t* tests and Wilcoxon signed-rank tests were performed to examine the change in exercise volumes, exertional oxygen flow rates, and functional outcomes. A McNemar test was used to examine the change in the number of LTx candidates who reported being active on the RAPA. A *P* value of <.05 was considered statistically significant. Statistical analyses were performed using SAS University Edition (SAS Institute Inc).

Results

Overview

There were 108 total participants including 78 LTx candidates (including 3 who were also LTx recipients during the study period) and 30 participants who only used the app as LTx recipients. Between March 16 and August 1, 2020, 84 people were active on the LTx wait list, of which 78 used the app for at least four weeks by September 1, 2020. Reasons for exclusion included no smartphone or tablet (n=1), no cellular data alongside limited Wi-Fi (n=2), declined (n=1), inpatient (n=1), and underwent LTx in less than four weeks (n=1; Figure 1). Between February 1 and July 15, 2020, there were 45 LTx recipients who could have used the app for at least four weeks between hospital discharge and three months posttransplant during the study period, of which 33 were included. Reasons for exclusion included no smartphone or tablet (n=2), no cellular data alongside limited Wi-Fi (n=1), declined (n=3), died early posttransplant (n=1), and inpatient (n=5; Figure 1). Patient demographics are reported in Tables 1 and 2.

Figure 1. Flow and attrition of lung transplant candidates and recipients. [a]This time frame would permit at least 4 weeks of rehabilitation data to be entered into the app between March 16, 2020, and September 1, 2020, accounting for 2 weeks of hospitalization posttransplant.

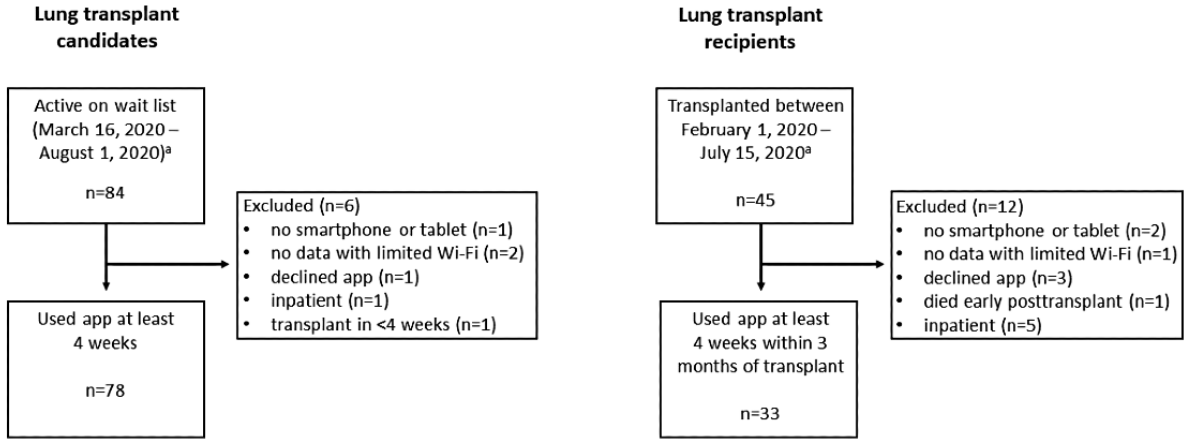


Table 1. Demographics of lung transplant candidates undergoing telerehabilitation (n=78)[a].

Characteristic	Values
Age (years), mean (SD)	59 (12)
Male sex, n (%)	37 (47)
Diagnosis, n (%)	
Interstitial lung disease	39 (50)
Chronic obstructive pulmonary disease	27 (35)
Cystic fibrosis	1 (1)
Pulmonary hypertension	5 (7)
Bronchiectasis	2 (2)
Re-transplant	4 (5)
Forced expiratory volume in one second (% predicted), mean (SD)	
Restrictive disease	52 (16)
Obstructive disease	26 (15)
Vascular disease	71 (17)
Forced vital capacity (% predicted), mean (SD)	
Restrictive disease	51 (17)
Obstructive disease	60 (12)
Vascular disease	85 (15)
Six-minute walk distance at transplant assessment (meters), mean (SD)	323 (109)
Six-minute walk distance at transplant assessment (% predicted) [19], mean (SD)	48 (16)
Short Physical Performance Battery at transplant assessment, median (IQR)	11 (9-12)
Fraction of inspired oxygen used during exercise [20], median (IQR)	0.4 (0.32-0.53)

[a]Used remote monitoring app for at least four weeks while listed for transplant between March 16, 2020, and September 1, 2020.

Table 2. Demographics of lung transplant recipients undergoing telerehabilitation (n=33)[a].

Characteristic	Values
Age (years), mean (SD)	58 (12)
Male sex, n (%)	20 (61)
Diagnosis, n (%)	
Interstitial lung disease	19 (58)
Chronic obstructive pulmonary disease	7 (21)
Cystic fibrosis	5 (15)
Pulmonary hypertension	2 (6)
Type of transplant, n (%)	
Double lung transplant	30 (91)
Single lung transplant	2 (6)
Double lung transplant-liver	1 (3)
Six-minute walk distance at 3 months (meters)[b], n=8, mean (SD)	422 (122)
Six-minute walk distance at 3 months (% predicted), n=8, mean (SD)	62 (18)
Forced expiratory volume in one second at 3 months (liters)[b], n=11, mean (SD)	2.3 (0.7)
Forced expiratory volume in one second at 3 months (% predicted), n=11, mean (SD)	73 (19)
Intensive care unit, length of stay (days), median (IQR)	4 (2-7)
Hospital length of stay (days), median (IQR)	21 (15-32)

[a]Used remote monitoring app for at least four weeks between hospital discharge and three months posttransplant between March 16, 2020, and September 1, 2020. Note: three patients were both transplant candidates and recipients during this period.

[b]Due to the COVID-19 restrictions, in-person 6-minute walk tests and pulmonary function tests performed at three months posttransplant were not routinely conducted between March 2020 and July 2020.

Satisfaction and Usage

Pretransplant, 42 LTx candidates and 21 LTx recipients completed the satisfaction survey administered in the app. Overall, 37 of 42 LTx candidates (88%) liked the virtual care features (videoconferencing, texting, education library, symptom surveys) and 35 of 42 (83%) agreed that it helped to prepare them for surgery (Multimedia Appendix 1). Posttransplant, 18 of 21 LTx recipients (85%) reported texting was helpful in self-recovery at home, which was higher than the number agreeing that videoconferencing, daily symptom check-ins, and educational health tips supported self-recovery (Multimedia Appendix 1). There was high usage by both patients and providers, including 365 video visits performed by the three program physiotherapists and widely accessed rehabilitation education materials in the app by patients (Multimedia Appendix 3). A total of 50 of 78 (64%) LTx candidates and 17 of 33 (51%) LTx recipients entered ≥10 prescribed exercise sessions into the app during the study time frame.

Physiotherapists reported overall satisfaction using the app to maintain communication, provide virtual support, and remotely monitor patients during the pandemic (n=3). There was agreement that the app supported ongoing access to patient educational resources, patient communication, and monitoring trends in exercise and biometric responses. Physiotherapists did not feel fully confident conducting remote clinical assessments using the app or identifying an early clinical change, and preferred to bring patients on site for functional or exertional oxygen reassessment when possible. The strongest barrier (rated as a 4 [barrier] or 5 [very strong barrier]) reported by all three physiotherapists was patient access to equipment and monitoring devices (Multimedia Appendix 2). An additional area listed as a barrier or very strong barrier by at least two of the physiotherapists included a lack of integration with Bluetooth devices for biometrics such as pulse oximeters, activity trackers, and exercise equipment. Physiotherapists preferred texting over traditional phone calls. Virtual visits were scheduled every 1-2 weeks, with 70% video to 30% phone visits.

Equipment Access

At the time of app registration, 48 of 78 (62%) LTx candidates completed a home equipment survey that was sent in the app, and 43 of 48 (90%) reported owning oximeters, although these were not necessarily medical grade (Multimedia Appendix 4). There was inconsistent access to exercise equipment at home. In addition, 17 of 48 (35%) reported being home alone during the day when exercising. Home equipment may have been purchased or obtained after this one-time survey was administered, and caregivers may have shifted to working from home, thus reducing the number of people who were alone during the day. Finally, 22 of 111 (20%) LTx candidates and recipients had hardware or software issues that impacted the app's videoconferencing feature, but they were still able to text and enter biometric data.

Exertional Oxygen Usage and Titration During Home Rehabilitation

An oxygen titration range was provided in the electronic medical record upon transplant listing after consultation between the respirologist and physiotherapist; patients are typically ordered to maintain an oxygen saturation of ≥88% with exercise. LTx candidates reported the following in the app: oxygen flow rate, oxygen delivery system, and oxygen source used. They also specified if they exercised on continuous versus pulsed oxygen delivery. Oxygen saturation, heart rate, and symptoms of dyspnea and fatigue were recorded after aerobic exercise. Oxygen flow ranged from room air to 20 liters per minute. A total of 58 of 78 (74%) LTx candidates increased their oxygen flow rate for aerobic exercise over time from 5 (IQR 3-10) liters per minute to 5.5 (IQR 3.5-15) liters per minute (P<.001).

Oxygen devices prescribed for home use included regular- and high-flow nasal cannulae, Oxymizer, Venturi mask, OxyMask, and non-rebreather mask. Upon the advice of physiotherapists, 13 of 66 (20%) LTx candidates reported changing their oxygen delivery device for aerobic exercise.

Surveys on Physical Activity and Exercise Self-efficacy

At the time of app registration, 13 of 23 (57%) LTx candidates self-reported as being active (eg, participating in 30 or more minutes of moderate intensity exercise 5 or more days per week) using the RAPA questionnaire, which improved to 20 of 23 (87%) after four weeks (P=.02; Figure 2). In addition, 37 of 78 (47%) LTx candidates completed the SEE at baseline and after four weeks. Depending on the individual, confidence for exercising regularly when alone increased (n=17, 46%), decreased (n=5, 14%), or remained the same (n=15, 40%). At three months posttransplant, 13 of 17 (76%) LTx recipients scored as active on the RAPA (Figure 3) and 12 of 33 (36%) completed the SEE, with a total mean SEE score of 74 (SD 11), indicating a high level of confidence that they could exercise under different conditions.

Figure 2. Number of lung transplant candidates who self-reported as being physically active on the Rapid Assessment of Physical Activity scale at baseline after app registration and after four weeks of home exercise (n=23). Scored as participating in 30 minutes or more of moderate intensity activity for 5 or more days per week.

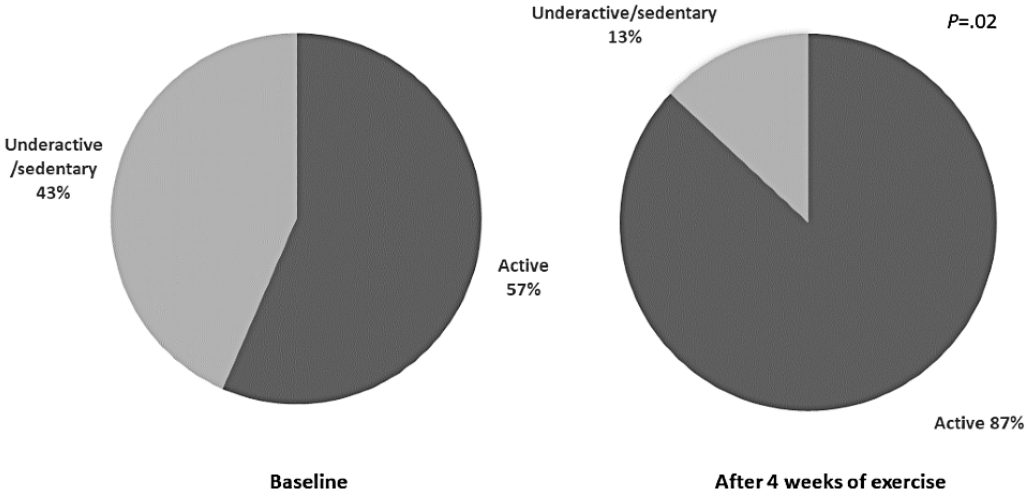
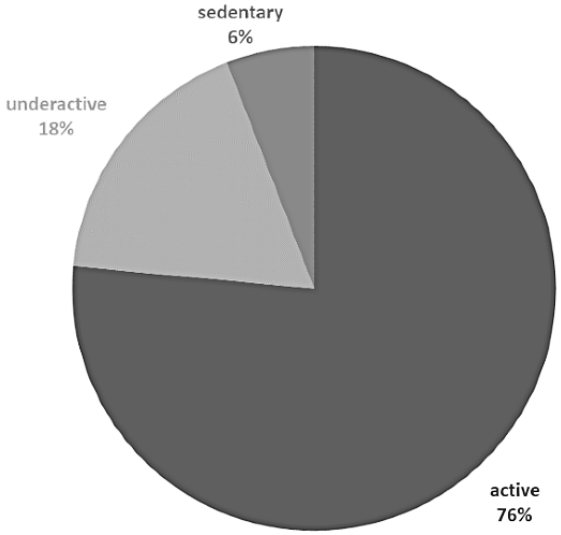


Figure 3. Categories of physical activity using the Rapid Assessment of Physical Activity scale three months posttransplant (n=17). Active: 30 minutes or more of moderate intensity physical activity 5 or more days per week. Underactive: some moderate physical activity but not every week or less than 30 minutes per day. Sedentary: rarely or never do any physical activities.



Exercise Volumes During Home Rehabilitation

Overall, 48 of 78 (62%) LTx candidates reported participating in non-treadmill walking exercise, which was recorded as steps (range 230-4847 steps), distance (18 meters to 3.2 kilometers), or time (3-80 minutes). In addition, 26 of 78 (33%) LTx candidates used a treadmill (range 0.5-2.8 mph) for 5-45 minutes. Over time, walking increased in duration (from 16 to 22 minutes; *P*=.002) but not speed (from 1.7 to 1.75 mph; *P*=.31; Table 3). A total of 37 of 78 (47%) LTx candidates had access to leg weights, and quadriceps weight used for leg extension did not change (3.5 [IQR 2.4-5] versus 4.3 [IQR 3-5] pounds; *P*=.08). Traditionally, 1 set of 10 repetitions was prescribed for center-based rehabilitation, with progression in the amount of weight lifted. During home rehabilitation, progression of weight was in part limited by access to equipment, as only 37 of 48 (75%) LTx candidates reported access to weights, which included primarily dumbbells for upper extremity training (Multimedia Appendix 4). Without the ability to increase the weight, increased sets were recommended to increase exercise training volume and 60 of 78 (77%) LTx candidates and 24 of 33 (73%) LTx recipients reported 2 or 3 sets of 10 repetitions for resistance training. In addition, 9 of 33 (27%) LTx recipients had access to a treadmill and increased treadmill speed (from 1.9 to 2.7 mph; *P*=.003) over a mean of 26 minutes (Table 3). Non-treadmill walking was recorded as time (range 11-90 minutes) and steps (1902-15,903 steps). Quadriceps weight increased (3 [IQR 0-3] versus 5 [IQR 3.8-6.5] pounds; *P*<.001; n=15).

Table 3. Changes to function and exercise training pretransplant and posttransplant after four or more weeks of home rehabilitation.

Outcome measures	Baseline	After ≥4 weeks of rehabilitation	*P* value
Lung transplant candidates			
Six-minute walk distance (meters), n=45, mean (SD)	346 (84)	307 (85)	.002
Total Short Physical Performance Battery, n=42, median (IQR)	12 (9.5-12)	12 (10-12)	.90
Treadmill speed (mph), n=26, mean (SD)	1.7 (0.6)	1.75 (0.6)	.31
Treadmill duration (minutes), n=26, mean (SD)	16 (9)	22 (10)	.002
Quadriceps weight (pounds), n=37[a], median (IQR)	3.5 (2.4-5)	4.3 (3-5)	.08
Lung transplant recipients			
Treadmill speed (mph), n=9, mean (SD)	1.9 (0.7)	2.7 (0.7)	.003
Treadmill duration (minutes), n=9, mean (SD)	19 (8)	26 (8)	.07
Quadriceps weight (pounds), n=15[a], median (IQR)	3 (0-3)	5 (3.8-6.5)	<.001

[a]Traditionally, 1 set of 10 repetitions is prescribed during our center-based rehabilitation, with progression in the amount of weight lifted. During home rehabilitation, progression of weight was limited by access to equipment and therefore increased sets were recommended.

Functional Outcomes

There was a decrease in pretransplant 6-minute walk distance (6MWD) from a mean of 346 (SD 84) meters to mean 307 (SD 85) meters (*P*=.002; n=45), and no change in the SPPB (12 [IQR 9.5-12] versus 12 [IQR 10-12]; *P*=.90; n=42). The 6MWT was performed in-person at the center, and the SPPB was performed either in-person or remotely with video supervision. Due to COVID-19 restrictions with on-site assessments, only 8 LTx recipients underwent a 6MWT three months posttransplant (5 men, mean 59 [SD 8] years, 75% interstitial lung disease). The mean 6MWD was 422 (SD 122) meters or 62% (SD 18%) predicted.

Discussion

During the COVID-19 pandemic, a rapid and large-scale clinical implementation of telerehabilitation for LTx candidates and recipients occurred that enabled exercise participation and progression. Despite the rapid implementation of a new model of care delivery and technology platform, usage and satisfaction were high. This early experience will guide program improvements and the development of an even more comprehensive and effective telerehabilitation program for the future.

Functional outcomes were lower compared to recent data of our center-based program where 6MWD was preserved during short-term prehabilitation and the SPPB improved pretransplant [20,21]. Improvements in pretransplant exercise volumes were lower with telerehabilitation than what has been seen historically in our center-based rehabilitation program (ie, increased treadmill speed and quadriceps weight used for resistance training) [4,19,20]. Multiple factors may have contributed to this. First, we had just initiated a small clinical project and had not received feedback from patients or providers to inform the co-design of optimal platform content, format, or delivery. In addition, clinical workflows, staffing, and technology support had not been mapped out for large scale implementation, and providers had to pivot their care model quickly due to COVID-19 restrictions with little to no experience in virtual care. Second, a pandemic environment increases barriers to exercise participation. Communal gym access was closed, and there was an increased demand and therefore long wait time to purchase home exercise equipment. People with chronic lung disease and those who are immunosuppressed were advised to socially distance and avoid leaving their homes for nonessential purposes [22]. In a study of 327 patients with cystic fibrosis (25% LTx recipients), 45% reported engaging in less physical activity during a lockdown between March 16 to May 16, 2020 [23]. Third, the lack of on-site exercise assessments and limited

evidence for remote functional assessments [16] may have led to an underprescription of exercise intensity. The ability to assess the degree of oxygen desaturation with medical-grade oximetry and closely supervise LTx candidates on-site was reduced, and extra vigilance with safety may have reduced the recommended exercise intensity, duration, and volume, and thus the efficacy of prehabilitation. Quantifying walking speed and progression for patients who did not have a treadmill and were walking inside their homes was more challenging. In our center-based program, LTx candidates could be switched to a longer cycling session and/or arm ergometry for aerobic training if they were not able to maintain adequate oxygenation on the treadmill or with hall-walking, and this was not always an option remotely.

A recent position statement from the Canadian Thoracic Society recommends caution when considering home or virtual pulmonary rehabilitation for patients with pulmonary hypertension, LTx candidates, and/or those with high oxygen requirements due to the limitations of home monitoring and lack of data around optimal exercise prescription in an unsupervised environment [10]. However, it is important that LTx candidates (who often have high oxygen requirements and include patients with pulmonary hypertension) participate in exercise to increase fitness for surgery, as listing for LTx did not stop during the pandemic. Although not formally tracked in the app, there were no serious adverse effects reported to the physiotherapy team or recorded in an incident report. As only a small number of LTx recipients underwent a 6MWT three months after transplant, it is not clear if LTx recipients reach the same functional benefit exercising at home versus a center-based program of supervised exercise three days per week from hospital discharge to three months posttransplant. Historically, LTx recipients achieve a 6MWD between 64%-76% predicted three months posttransplant [1,21].

Another concern around telerehabilitation is health equity [24], as not all patients were able to use the app if neither they nor a caregiver owned a compatible smartphone or tablet, and they did not have cellular data or reliable access to Wi-Fi. For security reasons, the app underwent regular Zoom updates and people who owned a phone/tablet with a lower-level operating system would experience connectivity difficulties for videoconferencing. Future work will include exploration of equipment libraries for devices (pulse oximeters, activity trackers), access to Wi-Fi or cellular data, and access to home exercise equipment. Remote patient monitoring, if applied thoughtfully and equitably, could allow patients to safely and effectively participate in rehabilitation remotely, thereby reducing some unnecessary travel to the transplant center. This can allow providers to better focus and prioritize in-person resources for patients who require them (eg, high and/or increasing exertional oxygen requirements, disease progression/exacerbation and symptom escalation, low and/or declining functional capacity, poor adherence and/or motivation for unsupervised home exercise), while continuing to closely monitor patients for issues and progress their exercise programs.

There are several limitations related to the design and context of this study. This was a program evaluation of a single center that broadly implemented telerehabilitation by necessity for infection control to limit on-site visits during the first wave of the COVID-19 pandemic. Subsequently, this did not permit a comparison arm of patients who did not receive telerehabilitation during the same period. Although we compared results to a group of historical controls attending on-site rehabilitation, the pandemic environment presented unique challenges, and it is not clear to what extent our findings are a result of the telerehabilitation model or related to contextual challenges during the pandemic. Next steps to further increase the strength of the evidence base supporting telerehabilitation and remote patient monitoring in LTx candidates and recipients include studies comparing different models of care in a postpandemic environment.

Additional future directions include examining the efficacy of a hybrid rehabilitation model, validating remote functional assessments, ensuring that the development and delivery of a telerehabilitation model is grounded in health behavior change theories [25], further exploring patient perceptions of home-based exercise monitoring [26,27]; integrating automatic download of Bluetooth exercise equipment, serial oximetry, and activity trackers into a virtual clinical care platform [28,29]; and customizing remote monitoring to meet the unique needs of a heterogeneous LTx population. The use of telerehabilitation and remote monitoring to support physical activity beyond the early posttransplant period may mitigate the well-documented risks of developing or worsening cardiometabolic disease following LTx. Telerehabilitation may also be beneficial for other chronic lung diseases and other solid-organ transplant populations.

In conclusion, our program was able to deliver telerehabilitation to LTx candidates and recipients despite challenges around equipment access and reduced on-site functional assessment. This early experience will inform the development of a robust and equitable telerehabilitation model during the COVID-19 pandemic and beyond.

Acknowledgments

We gratefully acknowledge clinical project funding by the University Health Network's Sprott Department of Surgery, the Toronto Lung Transplant Program, the Faire and Fecan Rehabilitation Fund, and the Soham & Shaila Ajmera Transplant Program. The authors acknowledge project management support from Greg Zilinskas and Caterina Masino, UHN Connected Care.

Conflicts of Interest

DR receives financial support from the Sandra Faire and Ivan Fecan Professorship in Rehabilitation Medicine. The other authors have no conflicts to declare.

Multimedia Appendix 1
Pretransplant app patient satisfaction survey.
[DOCX File , 24 KB - mhealth_v9i6e28708_app1.docx]

Multimedia Appendix 2
Physiotherapist app satisfaction survey.
[DOCX File , 25 KB - mhealth_v9i6e28708_app2.docx]

Multimedia Appendix 3
Patient and healthcare provider app usage (March 16-September 1, 2020).
[DOCX File , 24 KB - mhealth_v9i6e28708_app3.docx]

Multimedia Appendix 4
Pretransplant baseline rehabilitation survey (n=48).
[DOCX File , 24 KB - mhealth_v9i6e28708_app4.docx]

References

1. Wickerson L, Mathur S, Singer LG, Brooks D. Physical activity levels early after lung transplantation. Phys Ther 2015 Apr;95(4):517-525. [doi: 10.2522/ptj.20140173] [Medline: 25504488]
2. Langer D, Burtin C, Schepers L, Ivanova A, Verleden G, Decramer M, et al. Exercise training after lung transplantation improves participation in daily activity: a randomized controlled trial. Am J Transplant 2012 Jun;12(6):1584-1592 [FREE Full text] [doi: 10.1111/j.1600-6143.2012.04000.x] [Medline: 22390625]
3. Walsh JR, Chambers DC, Davis RJ, Morris NR, Seale HE, Yerkovich ST, et al. Impaired exercise capacity after lung transplantation is related to delayed recovery of muscle strength. Clin Transplant 2013;27(4):E504-E511. [doi: 10.1111/ctr.12163] [Medline: 23815281]
4. Li M, Mathur S, Chowdhury NA, Helm D, Singer LG. Pulmonary rehabilitation in lung transplant candidates. J Heart Lung Transplant 2013 Jun;32(6):626-632. [doi: 10.1016/j.healun.2013.04.002] [Medline: 23701852]
5. Martinu T, Babyak MA, O'Connell CF, Carney RM, Trulock EP, Davis RD, INSPIRE Investigators. Baseline 6-min walk distance predicts survival in lung transplant candidates. Am J Transplant 2008 Jul;8(7):1498-1505 [FREE Full text] [doi: 10.1111/j.1600-6143.2008.02264.x] [Medline: 18510641]
6. Armstrong HF, Garber CE, Bartels MN. Exercise testing parameters associated with post lung transplant mortality. Respir Physiol Neurobiol 2012 Apr 30;181(2):118-122. [doi: 10.1016/j.resp.2012.02.003] [Medline: 22503816]
7. Didsbury M, McGee RG, Tong A, Craig JC, Chapman JR, Chadban S, et al. Exercise training in solid organ transplant recipients: a systematic review and meta-analysis. Transplantation 2013 Mar 15;95(5):679-687. [doi: 10.1097/TP.0b013e31827a3d3e] [Medline: 23364480]
8. Wickerson L, Rozenberg D, Janaudis-Ferreira T, Deliva R, Lo V, Beauchamp G, et al. Physical rehabilitation for lung transplant candidates and recipients: An evidence-informed clinical approach. World J Transplant 2016 Sep 24;6(3):517-531 [FREE Full text] [doi: 10.5500/wjt.v6.i3.517] [Medline: 27683630]
9. Guidance for re-opening pulmonary rehabilitation programs. American Thoracic Society Assembly on Pulmonary Rehabilitation. URL: https://www.thoracic.org/members/assemblies/assemblies/pr/resources/ats-pr-assembly-re-opening-pr-document-final.pdf [accessed 2020-10-23]
10. Dechman G, Aceron R, Beauchamp M, Bhutani M, Bourbeau J, Brooks D, et al. Delivering pulmonary rehabilitation during the COVID-19 pandemic: A Canadian Thoracic Society position statement. Canadian Journal of Respiratory, Critical Care, and Sleep Medicine 2020 Nov 23;4(4):232-235. [doi: 10.1080/24745332.2020.1828683]
11. Garvey C, Holland AE, Corn J. Pulmonary rehabilitation resources in a complex and rapidly changing world. American Thoracic Society. URL: https://www.thoracic.org/members/assemblies/assemblies/pr/resources/pr-resources-in-a-complex-and-rapidly-changing-world-3-27-2020.pdf [accessed 2020-10-23]
12. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: Review of the State-of-the-Art and Areas of Application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7 [FREE Full text] [doi: 10.2196/rehab.7511] [Medline: 28733271]
13. Choi J, Hergenroeder AL, Burke L, Dabbs AD, Morrell M, Saptono A, et al. Delivering an in-Home Exercise Program via Telerehabilitation: A Pilot Study of Lung Transplant Go (LTGO). Int J Telerehabil 2016;8(2):15-26 [FREE Full text] [doi: 10.5195/ijt.2016.6201] [Medline: 28775798]
14. Choi J, Hergenroder A, Kostra J. Delivering a home-based telerehabilitation exercise program for a lung transplant candidate. In: Am J Respir Crit Care Med. 2017 Presented at: ATS 2017; May 2017; Washington, DC p. A2341 URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2017.195.1_MeetingAbstracts.A2341

15. Singer JP, Soong A, Bruun A, Bracha A, Chin G, Hays SR, et al. A mobile health technology enabled home-based intervention to treat frailty in adult lung transplant candidates: A pilot study. Clin Transplant 2018 Jun;32(6):e13274 [FREE Full text] [doi: 10.1111/ctr.13274] [Medline: 29742287]
16. Holland AE, Malaguti C, Hoffman M, Lahham A, Burge AT, Dowman L, et al. Home-based or remote exercise testing in chronic respiratory disease, during the COVID-19 pandemic and beyond: A rapid review. Chron Respir Dis 2020;17:1479973120952418 [FREE Full text] [doi: 10.1177/1479973120952418] [Medline: 32840385]
17. Topolski TD, LoGerfo J, Patrick DL, Williams B, Walwick J, Patrick MB. The Rapid Assessment of Physical Activity (RAPA) among older adults. Prev Chronic Dis 2006 Oct;3(4):A118 [FREE Full text] [Medline: 16978493]
18. Resnick B, Jenkins LS. Testing the reliability and validity of the Self-Efficacy for Exercise scale. Nurs Res 2000;49(3):154-159. [doi: 10.1097/00006199-200005000-00007] [Medline: 10882320]
19. Hill K, Wickerson LM, Woon LJ, Abady AH, Overend TJ, Goldstein RS, et al. The 6-min walk test: responses in healthy Canadians aged 45 to 85 years. Appl Physiol Nutr Metab 2011 Oct;36(5):643-649. [doi: 10.1139/h11-075] [Medline: 21967531]
20. Wickerson L, Brooks D, Reid WD, Singer LG, Granton J, Mathur S. Exertional Oxygen Requirements During Exercise Training in Advanced Interstitial Lung Disease. J Cardiopulm Rehabil Prev 2018 Nov;38(6):419-424. [doi: 10.1097/HCR.0000000000000338] [Medline: 29952810]
21. Wickerson L, Rozenberg D, Gottesman C, Helm D, Mathur S, Singer LG. Pre-transplant short physical performance battery: Response to pre-habilitation and relationship to pre- and early post-lung-transplant outcomes. Clin Transplant 2020 Dec;34(12):e14095. [doi: 10.1111/ctr.14095] [Medline: 32970883]
22. Reduce your risk from COVID-19. Public Health Ontario. URL: https://tinyurl.com/mrcd2vca [accessed 2020-12-11]
23. Radtke T, Haile SR, Dressel H, Benden C. Recommended shielding against COVID-19 impacts physical activity levels in adults with cystic fibrosis. J Cyst Fibros 2020 Nov;19(6):875-879 [FREE Full text] [doi: 10.1016/j.jcf.2020.08.013] [Medline: 32878732]
24. 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] [Medline: 32452816]
25. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
26. Dobson R, Herbst P, Candy S, Brott T, Garrett J, Humphrey G, et al. Understanding End-User Perspectives of Mobile Pulmonary Rehabilitation (mPR): Cross-Sectional Survey and Interviews. JMIR Form Res 2019 Dec 20;3(4):e15466 [FREE Full text] [doi: 10.2196/15466] [Medline: 31859681]
27. Argent R, Daly A, Caulfield B. Patient Involvement With Home-Based Exercise Programs: Can Connected Health Interventions Influence Adherence? JMIR mHealth uHealth 2018 Mar 01;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
28. Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, et al. Wearable Health Devices in Health Care: Narrative Systematic Review. JMIR mHealth uHealth 2020 Nov 09;8(11):e18907 [FREE Full text] [doi: 10.2196/18907] [Medline: 33164904]
29. Gay V, Leijdekkers P. Bringing Health and Fitness Data Together for Connected Health Care: Mobile Apps as Enablers of Interoperability. J Med Internet Res 2015 Nov 18;17(11):e260 [FREE Full text] [doi: 10.2196/jmir.5094] [Medline: 26581920]

Abbreviations

LTx: lung transplant
RAPA: Rapid Assessment of Physical Activity
SEE: Self-efficacy For Exercise
6MWD: 6-minute walk distance
6MWT: 6-minute walk test
SPPB: Short Physical Performance Battery

Edited by L Buis; submitted 11.03.21; peer-reviewed by A Saptono, K Moor; comments to author 11.04.21; revised version received 30.04.21; accepted 27.05.21; published 17.06.21.

Please cite as:
Wickerson L, Helm D, Gottesman C, Rozenberg D, Singer LG, Keshavjee S, Sidhu A
Telerehabilitation for Lung Transplant Candidates and Recipients During the COVID-19 Pandemic: Program Evaluation
JMIR Mhealth Uhealth 2021;9(6):e28708
URL: https://mhealth.jmir.org/2021/6/e28708
doi:10.2196/28708
PMID:34048354

Original Paper

European Portuguese Version of the User Satisfaction Evaluation Questionnaire (USEQ): Transcultural Adaptation and Validation Study

Célia Domingos[1,2,3,4], BSc; Patrício Soares Costa[1,2,5], PhD; Nadine Correia Santos[1,2,6], PhD; José Miguel Pêgo[1,2,3,4], MD, PhD

[1]Life and Health Sciences Research Institute, School of Medicine, University of Minho, Braga, Portugal
[2]ICVS/3B's, PT Government Associate Laboratory, Braga, Portugal
[3]iCognitus4ALL - IT Solutions, Braga, Portugal
[4]Clinical Academic Center - Braga, Braga, Portugal
[5]Faculty of Psychology and Education Sciences, University of Porto, Braga, Portugal
[6]Associação Centro de Medicina Digital P5, Braga, Portugal

Corresponding Author:
José Miguel Pêgo, MD, PhD
Life and Health Sciences Research Institute
School of Medicine
University of Minho
Campus de Gualtar
Braga, 4710-057
Portugal
Phone: 351 253604932
Email: jmpego@med.uminho.pt

Abstract

Background: Wearable activity trackers have the potential to encourage users to adopt healthier lifestyles by tracking daily health information. However, usability is a critical factor in technology adoption. Older adults may be more resistant to accepting novel technologies. Understanding the difficulties that older adults face when using activity trackers may be useful for implementing strategies to promote their use.

Objective: The purpose of this study was to conduct a transcultural adaptation of the User Satisfaction Evaluation Questionnaire (USEQ) into European Portuguese and validate the adapted questionnaire. Additionally, we aimed to provide information about older adults' satisfaction regarding the use of an activity tracker (Xiaomi Mi Band 2).

Methods: The USEQ was translated following internationally accepted guidelines. The psychometric evaluation of the final version of the translated USEQ was assessed based on structural validity using exploratory and confirmatory factor analyses. Construct validity was examined using divergent and discriminant validity analysis, and internal consistency was evaluated using Cronbach α and McDonald ω coefficients.

Results: A total of 110 older adults completed the questionnaire. Confirmatory factor analysis supported the conceptual unidimensionality of the USEQ (χ^2_4=7.313, *P*=.12, comparative fit index=0.973, Tucker-Lewis index=0.931, goodness of fit index=0.977, root mean square error of approximation=0.087, standardized root mean square residual=0.038). The internal consistency showed acceptable reliability (Cronbach α=.677, McDonald ω=0.722). Overall, 90% of the participants reported excellent satisfaction with the Xiaomi Mi Band 2.

Conclusions: The findings support the use of this translated USEQ as a valid and reliable tool for measuring user satisfaction with wearable activity trackers in older adults, with psychometric properties consistent with the original version.

(JMIR Mhealth Uhealth 2021;9(6):e19245) doi:10.2196/19245

KEYWORDS

satisfaction; usability; reliability; validity; seniors; elderly; technology; wearables

Introduction

The use of mobile health (mHealth) technology has greatly increased over the past decade. These technologies enable health promotion and self-monitoring of health-related behaviors [1-3] and show potential in disease treatment and prevention in a cost-efficient and widely accessible manner [4]. Currently, mHealth covers a wide range of technologies, such as wearable devices and smartphone apps, for tracking different types of health-related data including physical activity [4,5].

Activity trackers are sensor-based wearable devices that automatically track and monitor various indicators of physical activity (eg, steps, calories burned, and distance traveled), with some also able to record heart rate and sleep measures [2,6-8]. The technology has the potential to help older adults with health self-management and self-efficacy by improving lifestyle behaviors and motivating compliance or attainment of daily activity goals [3,6,9]. Despite this potential, most older adults do not use activity trackers for their health-tracking needs. A possible explanation may be a matter of usability. Although it is a critical factor that can determine technology adoption, these devices have been mainly developed for a younger target group; thus, older adults may have difficulties due to usability barriers [2,9].

Published by the International Organization for Standardization, ISO-9241-11 defines usability in terms of effectiveness, efficiency, and user satisfaction rating of a product in a specific environment, by a specific user, for a specific purpose [10,11]. Accordingly, user satisfaction can be thought of as a usability component, but it cannot be evaluated in the same manner as efficiency and effectiveness. Satisfaction is the user's attitude toward the system they use, affecting behavior intention for continuous or future use [12-14]. Moreover, satisfaction depends on how comfortable the user feels using the system [10,11,14,15]. Despite the importance of user satisfaction to ensure usability and the improved development of mHealth solutions, there is a gap in research assessing user satisfaction with mHealth [4].

Currently, there are a few validated and widely used questionnaires to collect targeted user feedback for the evaluation of a system's usability. These include the System Usability Scale (SUS) [15], the Post-Study System Usability Questionnaire (PSSUQ) [10,16-18], and the Usefulness, Satisfaction and Ease of Use (USE) questionnaire [10,17,19]. Of these, only the SUS [20] and the USE [21] questionnaires are available in European Portuguese for generalized evaluation of usability. Regarding user satisfaction questionnaires, a study by Melin et al [4] presented the development of the mHealth Satisfaction Questionnaire, the Questionnaire for User Interaction Satisfaction (QUIS) [22], and the User Satisfaction Evaluation Questionnaire (USEQ). Nonetheless, despite this, there is no validated questionnaire for measuring user satisfaction with technologies available in European Portuguese.

Therefore, this study aimed to provide a valid questionnaire to specifically evaluate user satisfaction with technologies in Portuguese older adults. For this, the USEQ was selected since it is a short but comprehensive questionnaire with a reasonable number of questions; importantly, it is clear and easy to understand [23]. We also aimed to evaluate user satisfaction with the Xiaomi Mi Band 2 among older adults. Regardless of spoken language, and despite a growing number of studies on subjective experiences of user satisfaction with and the usability and usefulness of wearable technologies, only a few studies have focused on older adults [3,8,24]. Understanding whether a cohort of older adults is satisfied with these activity trackers is important to ensure devices can be successfully implemented in clinical and research settings [25].

Methods

The study was divided into two phases. Phase 1 addressed the translation of the USEQ questionnaire into European Portuguese and its cross-cultural adaptation. Phase 2 involved the assessment of the USEQ psychometric properties and its validation in the new context.

Translation and Cultural Adaptation of the USEQ

The USEQ is composed of 6 questions and uses a 5-point Likert scale for responses. The total score ranges from 6 (poor satisfaction) to 30 (excellent satisfaction). All questions are affirmative, except question 5, which is a negatively posed question. The numerical value of the affirmative questions is used to calculate the score. The negative question subtracts the numerical value of the response from 6 and then adds this result to the total score [23]. Since the USEQ was designed to evaluate user satisfaction with virtual rehabilitation systems, the questionnaire comprises one item that specifically measures the perceived usefulness of using the technology for rehabilitation. Thus, in this study, we adapted this item to include an item that could be applied to general-purpose systems or across different types of mHealth.

The adapted English version of the USEQ was culturally and linguistically adapted to European Portuguese after obtaining formal authorization from the original author (Gil-Gómez [23]). The process of translation followed the general guidelines provided by Lenz et al [26] and the World Health Organization [27]. Briefly, it comprised the following steps: forward translation, translation review and reconciliation of content, back translation, preliminary version, pretesting, and final version (Multimedia Appendix 1). In step 1 (forward translation), the USEQ questionnaire was translated to European Portuguese by two independent English-proficient translators, whose native language is European Portuguese. In step 2 (translation review and reconciliation of content), the independent translations were reviewed by both forward translators as well as by an independent team member. A reconciliation version of the document was obtained. In step 3 (back translation), the reconciled version was translated from European Portuguese into English by independent translators fluent in both languages, who were blinded to the original USEQ version of the document. The retroversion was done as a quality control step and to verify that both versions were equivalent. In step 4 (preliminary version), all team members performed a comparative analysis between the back-translation and original version. A preliminary version was prepared after items were reviewed and a consensus was reached. Finally, in step 5

(pretesting), the preliminary version of the document was tested in a pilot study with a sample of 20 independent and representative individuals selected from those who will be administered the final questionnaire. The individuals were asked to provide feedback on their understanding of the questions, the preferred use of alternative words for a given expression, terms deemed unacceptable or offensive, and their opinion of the questionnaire. The information collected was used to improve and develop the final version of the USEQ.

Psychometric Properties and Validation of the USEQ

The psychometric validation of the final version of the USEQ was based on real data collection after the users' experience with the wearable activity tracker. The analyses included assessment of structural validity, construct validity, and internal consistency.

To provide evidence of construct validity, the participants' responses on the USEQ were correlated with the pre-existing instruments that measure similar concepts—the SUS and the technology acceptance model 3 (TAM3; convergent validity). For divergent validity, the participants' responses on the USEQ were correlated with Mini-Mental State Examination (MMSE) scores [28]. Briefly, the SUS is the most widely used standardized questionnaire to measure perceived usability [10,15,29,30], while the TAM is the most applied theoretical model for evaluating or predicting users' acceptance of new technologies [31]. Lastly, the MMSE is a widely accepted questionnaire to assess cognitive function in older adults [32].

Participants and Data Collection

Following the application of inclusion and exclusion criteria, a total of 120 community-dwelling older adults (aged 64-75 years) from Northern Portugal were recruited to the study. The primary exclusion criteria were an inability to understand informed consent and neuropsychiatric and neurodegenerative disorders. Among the recruited participants, a final sample of 110 participants completed the usability test of the wearable activity tracker.

A baseline characterization was performed through a sociodemographic questionnaire and a standardized clinical interview. Moreover, since individual differences, including demographics, cognitive state, and emotional state influence individuals' perceptions regarding the technology [33], we included a neuropsychological evaluation to obtain mood (Geriatric Depression Scale [GDS]) [34] and global cognitive profiles (MMSE) [35].

For usability testing, the Xiaomi Mi Band 2 was selected among several commercially available wearable activity trackers, since it is ergonomic, accessible, easy to operate, and offers the best price-quality ratio. The device provides general health monitoring, combining sensors that allow objective assessment of activity levels, heart rate, and sleep patterns [5,36,37]. The participants used a Xiaomi Mi Band 2 over 15 days while performing their normal daily activities. They were instructed to wear the activity tracker continuously. After concluding the usage testing period, participants were asked to provide information about their experience. This was attained through application of the USEQ [23] to evaluate user satisfaction, the TAM 3 [38] to collect information about technology acceptance, and the SUS for perceived usability [20].

Statistical Analysis

Data analysis was performed using IBM SPSS Statistics (version 26; IBM Corp), and JASP (version 0.11.1). Descriptive statistics (mean, median, standard deviation, minimum, maximum, skewness, and kurtosis) were calculated for each variable. Normality was considered adequate if absolute values for skewness and kurtosis were above 2.0 and 7.0, respectively [39,40].

Structural Validity

The creators of the original scale analyzed the factor structure through principal component analysis (PCA) [23]. Results indicated two components with an eigenvalue greater than 1. The first component had all six items and explained 43% of the variance, and the second component had only four items (items 1, 4, 5, and 6), only two of which had factor loadings greater than 0.5. Therefore, after the analysis of the scree plot, they considered a one-factor solution explaining 42.9% of the variance to be appropriate.

Before conducting exploratory factor analysis (EFA) with our data, the "parameters" R package was used to decide the number of factors to extract. The solution for one dimension was supported by 6 (42.9%) methods of 14 (acceleration factor, standard error scree, Tucker-Lewis index [TLI], root mean square error of approximation [RMSEA], adjusted root mean square residual, and Bayesian information criterion) [41]. As PCA is a data reduction technique that does not conduct to a latent variable model, principal axis factoring was used to extract the latent factor.

Variables with factor loadings above 0.4 were extracted. Before the analysis, the appropriateness of the data for factor analysis was examined using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett test of sphericity [42,43]. Regarding the sample size for the EFA, there is no consensus on the number of participants required to perform the analysis [44,45]. Hatcher et al [46] recommend a minimum subject to item ratio of at least 5:1, with a minimum of 100 subjects. The number of components was determined by Kaiser criteria (retaining factors with eigenvalues greater than 1), scree plot inspection [43], and parallel analysis [47].

Subsequently, confirmatory factor analysis (CFA) was performed using maximum likelihood estimation. The factor loadings were used as local indices of goodness of fit as well as the following goodness-of-fit indices and thresholds for a good fit: chi-square test (χ^2; P>.05), chi-square value divided by degrees of freedom (a ratio of ≤3), comparative fit index (CFI≥0.90), Tucker-Lewis index (TLI≥0.90), goodness of fit index (GFI ≥0.90), root mean square error of approximation (RMSEA<0.08) and standardized root mean square residual (SRMR≤0.08) [48-50].

Ideally, a CFA should be performed in a subsequent study using another sample (validation sample) [51]. However, due to having an insufficient sample size to perform the analysis in two

subsamples, the same sample was used for both approaches (EFA and CFA).

Internal Consistency

Internal consistency was assessed using both Cronbach α and McDonald ω coefficients [52,53]. Cronbach α is the most widely used measure of reliability. However, it overestimates the true composite reliability and is negatively biased when used to measure the reliability of ordinal variables [52,54]. Given the ordinal response format of USEQ items, McDonald ω was used to overcome these limitations, providing a more accurate approximation of the scale reliability [55]. Cronbach α and McDonald ω coefficients over .70 are considered indicators of satisfactory item homogeneity [55,56]. Item-total and inter-item correlations were analyzed, considering cutoff values over .30 and under .70, respectively [57].

Construct Validity: Convergent and Divergent Validity

Convergent and divergent validity were estimated using Spearman correlation (*r*). It was hypothesized a priori that the USEQ score would be positively correlated with the SUS score and TAM 3 score, while correlation was not expected with the MMSE score. A coefficient of *r*=0.3 is assumed to provide evidence of convergent and divergent validity [58,59].

USEQ Scores

Descriptive statistics and normality were assessed for the USEQ's total score. Correlations between USEQ scores and demographic as well as mood and global cognitive characteristics were estimated using Spearman correlation.

Ethics Statement

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local and national ethics committees (approval number 42-2018). The study goals and assessments were explained to potential participants. All participants provided written informed consent before study enrollment.

Results

Study Participants

A total of 110 participants completed the USEQ questionnaire. Demographic, mood, and global cognitive characteristics of the sample are presented in Table 1. Participants had an average age of 68.41 (SD 3.11) years with a mean of 7.95 (SD 5.38) years of education.

Table 1. Characteristics of the study participants (N=110).

Characteristics	Values
Age, years, mean (SD)	68.41 (3.11)
Gender, male, n (%)	50 (45.5)
Education, years of formal schooling, mean (SD)	7.95 (5.38)
Mini-Mental State Examination, total score, mean (SD)	26.95 (2.00)
Geriatric Depression Scale, total score, mean (SD)	6.05 (4.58)

Study participants responded to all items of the USEQ. Descriptive statistics for USEQ items are presented in Table 2. Given the ordinal nature of the variables assessed, item distribution demonstrates some degree of nonnormality. In fact, most data collected in behavioral research does not follow univariate normal distributions [40,60]. For the USEQ total score, the results reveal acceptable values for both skewness (sk=–2.02) and kurtosis (k=3.96), showing no severe violation of normality. Regarding the individual items, the kurtosis value was not acceptable for item 1; thus, it was excluded from further analysis.

Table 2. Descriptive statistics for User Satisfaction Evaluation Questionnaire items.

Items	Minimum	Maximum	Median	Mean (SD)	Skewness	Kurtosis
Item 1	1	5	5	4.80 (0.59)	–3.81	17.67
Item 2	3	5	5	4.82 (0.47)	–2.66	6.46
Item 3	2	5	5	4.65 (0.71)	–2.21	4.50
Item 4	2	5	5	4.47 (0.75)	–1.30	0.98
Item 5	1	5	5	4.65 (0.93)	–2.71	6.15
Item 6	1	5	5	4.70 (0.69)	2.68	8.30

Structural Validity

An exploratory factor analysis was conducted on the 5 items of the USEQ. The Kaiser-Meyer-Olkin measure demonstrated adequacy for the analysis (KMO=0.629) as a value above 0.6 indicates an adequate sample size [43]. The Bartlett sphericity test (χ^2_{10}=126, *P*<.001) indicated that the correlation between the items is sufficient to perform the analysis. The analysis showed that the one-factor solution explains 47% of the variance (Table 3) and comprises all items with factor loadings higher than 0.3. Our findings corroborated the decision of the original

questionnaire authors, who considered a one-factor solution to be appropriate.

Table 3. Factor matrix containing obliquely unrotated factor loadings of principal axis factoring (forcing one-factor solution). The eigenvalue and the percentage of variance explained by the factor are also shown.

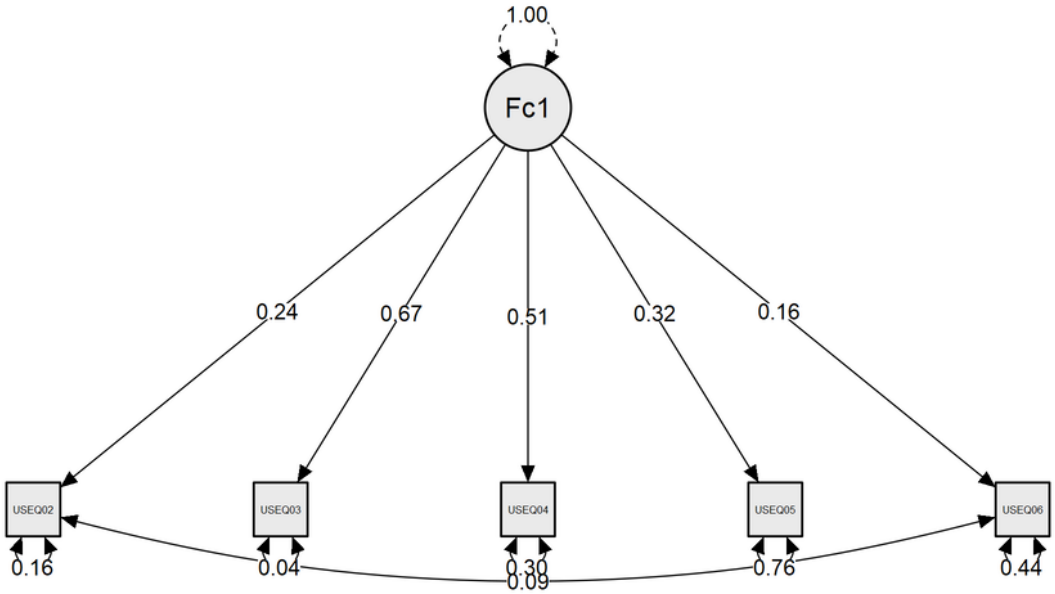
Items	Factor 1
Item 2	0.568
Item 3	0.870
Item 4	0.680
Item 5	0.403
Item 6	0.346
Eigenvalue	2.345
Percent of variance	46.9

A one-factor solution CFA model was tested for the interference dimension, as hypothesized by the original authors [23]. All 5 items were loaded onto a single latent variable. Table 4 shows goodness-of-fit measures for the model, revealing acceptable measures for the following indices: χ^2_4=7.313, *P*=.12; CFI=0.973, TLI=0.931, GFI=0.977, RMSEA=0.087, and SRMR=0.038 [61]. The final model is presented in Figure 1.

Table 4. Fit indices for confirmatory factor analysis model.

Indices	Model
Chi-square value (df)	7.313 (4)
Chi-square value to df ratio	1.83
P value	.12
Comparative fit index	0.973
Tucker-Lewis index	0.931
Goodness of fit index	0.977
Root mean square error of approximation	0.087
Standardized root mean square residual	0.038

Figure 1. Path diagram and standardized estimates for the one-factor model of the User Satisfaction Evaluation Questionnaire.



Internal Consistency

Analysis of the internal consistency showed acceptable reliability (Cronbach α=.677; McDonald ω=0.722), indicating a reliability homogeneity of the items for the one-factor solution. The corrected item-total correlation values ranged from 0.080 to 0.654, showing an adequate correlation of each item and suggesting adequate scale homogeneity. The inter-item correlations were all below 0.70, indicating nonredundancy of items (Table 5).

Table 5. Inter-item correlation matrix.

Items	Item 2	Item 3	Item 4	Item 5
Item 3	0.495	—[a]	—	—
Item 4	0.270	0.654	—	—
Item 5	0.251	0.317	0.327	—
Item 6	0.397	0.219	0.207	0.080

[a]Not applicable.

Convergent and Divergent Validity

The association between the USEQ and SUS was moderate in magnitude (r=0.43, P<.001), and strong between TAM 3 factor "perceived ease of use" (r=0.690, P<.001) and TAM 3 factor "perceptions of external control" (r=0.571, P<.001). No significant correlation was observed between USEQ and MMSE (r=0.042, P=.661).

USEQ Scores

To rate the USEQ score, the following classification was used: poor (0-5), fair (5-10), good (10-15), very good (15-20), or excellent (20-25) [62]. The mean USEQ score was 23.30 (SD 2.40), indicating an excellent level of satisfaction with the activity tracker (Xiaomi Mi Band 2; Table 6).

Table 6. Scores obtained on the User Satisfaction Evaluation Questionnaire for the original scale with 6 items and the newly proposed scale with 5 items.

Items	Minimum	Maximum	Mean (SD)	Skewness	Kurtosis
USEQ (5 items)	14	25	23.30 (2.40)	–1.81	2.99
USEQ (original)	17	30	28.07 (2.84)	–2.01	3.96

All participants reported user satisfaction experiences above "good," with 90% of all participants reporting excellent satisfaction with the device. Furthermore, a significant correlation (r=–0.319, P=.001) between depressive mood and user satisfaction was noted; a higher score on the GDS (ie, higher depressive mood) was negatively associated with satisfaction with the device (Table 7).

Table 7. Spearman bivariate correlations between User Satisfaction Evaluation Questionnaire scores and demographic, mood, and global cognitive characteristics.

Values	Gender	Age	Years of education	Mini-Mental State Examination	Geriatric Depression Scale
Spearman correlation coefficient (r)	–0.005	–0.120	0.124	0.042	–0.319[a]
P value	.96	.21	.20	.66	.001

[a]Significant correlation.

Discussion

Physical activity is associated with health benefits, a decreased burden of disease, and a decrease in all-cause mortality in adults [2]. Nowadays, wearable activity trackers provide the opportunity to increase physical activity levels through continuous monitoring [8], which may be especially beneficial for older adults. However, over 75% of the over-65 age group state that they require assistance to use new technologies [2]. Usability studies on wearable activity trackers are needed to better understand the barriers that older adults face when using these technologies. In a cohort of older adults, this study aimed to provide a valid questionnaire to evaluate user satisfaction using an activity tracker (Xiaomi Mi Band 2).

The results from the translation phase show that the items were easy to understand and that there were no semantic problems. Moreover, the translated items were considered equivalent to the original version. In particular, the USEQ was found to be a simple and easy-to-understand questionnaire with an appropriate number of questions to apply in older populations. Validity evidence was obtained with 5 questions, maintaining the original one-factor structure. Similar to the original study by Gil-Gómez et al [23], reporting that the one-factor solution explained about 43% of the variance, here an approximate 47% of the variance was explained by the one-factor solution.

Confirmatory factor analysis showed acceptable fit indexes (χ^2_4=7.313, P=.12, CFI=0.973, TLI=0.931, GFI=0.977, RMSEA=0.087, and SRMR=0.038). However, it is

recommended to evaluate structural validity in another independent study. The reliability results show that the European Portuguese version of the USEQ yielded an adequate internal consistency (Cronbach α=.677; McDonald ω=0.722) in a sample of older individuals. Overall, results indicate that the psychometric properties of the European Portuguese version of the USEQ are comparable with those of the original version, and therefore may be used for the evaluation of satisfaction concerning other technologies, including wearable activity trackers.

Convergent validity is one of the fundamental aspects of construct validity and it refers to how closely the new questionnaire is related to other variables and other measures of the same construct. Regarding convergent validity, as expected, the USEQ correlated with the SUS (r=0.43, P<.001). The SUS is a widely used standardized questionnaire for the assessment of the perceived usability of technology. Thus, a moderate positive correlation was expected, given that the two questionnaires are meant to measure similar constructs. TAM 3 has been one of the most influential models regarding technology acceptance. It distinguishes two concepts influencing an individual's intention to use new technology: perceived ease of use and perceived usefulness. Therefore, we also expected a correlation between some constructs of the TAM 3 and the USEQ. Indeed, results showed a positive moderate correlation of the USEQ with the TAM 3 factors "perceived ease of use" (r=0.690, P<.001) and "perceptions of external control" (r=0.571, P<.001).

No demographic variables were found to be significantly correlated to user satisfaction with the device. However, a higher depressive mood, as evaluated by the GDS, was negatively associated with the satisfaction perceived by participants using the device. Similarly, a recent study investigating the impact of depressive symptoms on measures of web user experience found a significant association between depressive symptoms and subjective user experience [63]. These results indicate that mood may be a factor influencing technology usability and may warrant further guidance and/or targeted approaches in the use of these technologies by specific populations.

Regarding user satisfaction with the Xiaomi Mi Band 2, the device achieved a score of 23.30 (SD 2.40) in the USEQ, demonstrating an excellent reported level of satisfaction and thus suggesting suitability for older adults. Furthermore, results showed that item 4 ("Is the information provided by the system clear?"), which is related to the perceived ease of use, yielded the lowest score. This result may suggest that older adults could have difficulties in understanding the information provided by the activity tracker, which should be noted by manufacturers. The perceived ease of use refers to the degree to which a person perceives how easy it is to use the technology and is one of the primary factors that affect an individual's intention to use new technology [7,9,31]. This kind of difficulty is especially interesting considering the age of the participants enrolled in the study. Older adults tend to perceive technologies as difficult to use due to usability problems related to poor memory, decreased vision, and poor literacy [64], but this may not necessarily be the case for all older individuals. Thus, results should be interpreted with caution and future studies should include cohorts with different characteristics (for instance, higher school levels or those that have been [early] adopters of different types of technologies).

Concerning limitations, the study was conducted using a convenience sample; therefore, the participants may not represent the entire older population. Moreover, if the sample used in this study is more homogenous than the wider population on the common factors, this can lead to attenuation in correlations and can influence the strength or bias of correlations among variables [51]. Future studies should have a larger sample, and it would be beneficial to the study to maximize variance on measured variables relevant to the constructs of interest [51]. It would also be of value to evaluate the long-term use of this device and motivations for long-term use. Moreover, studies combining quantitative and qualitative methods, such as interviews, would also be valuable to explore older adults' perceptions and experiences, and to provide details about user behaviors, user needs, and specific problems that quantitative measures cannot address. Finally, further user satisfaction studies of older adults using activity trackers should include other devices. This would ensure that such devices can be effectively implemented in clinical and research settings to promote physical activity.

In conclusion, the European Portuguese version of the USEQ has adequate psychometric properties consistent with the original version, supporting its use as a valid and reliable tool for measuring user satisfaction in older adults. Furthermore, we adapted USEQ to a generic questionnaire for user satisfaction that can be used with several mHealth technologies, including smartphones, patient monitoring devices, tablets, mobile health apps, personal digital assistants, and other wireless devices. Finally, this study has contributed to the currently available and growing body of information on the usability of wearable technologies among older adults.

Acknowledgments

Financial support was provided by the European Regional Development Fund (FEDER) through the Operational Programme Competitiveness Factors – COMPETE and National Funds through FCT under the projects POCI-01-0145-FEDER-007038, UIDB/50026/2020, and UIDP/50026/2020, by the projects NORTE-01-0145-FEDER-000013 and NORTE-01-0145-FEDER-000023 (supported by the North Portugal Regional Operational Programme [NORTE 2020], under the Portugal 2020 [P2020] Partnership Agreement, through FEDER), by the project MEDPERSYST (POCI-01-0145-FEDER-016428; supported by the Operational Programme Competitiveness and Internationalization [COMPETE 2020] and the Regional Operational Program of Lisbon and National Funding through the Portuguese Foundation for Science and Technology [FCT, Portugal]), and by the Portuguese North Regional Operational Programme (ON.2 – O Novo Norte, under the National Strategic Reference Framework [QREN], through

FEDER). The work was also developed under the scope of the 2CA-Braga Grant of the 2017 Clinical Research Projects. CD was supported by a combined PhD scholarship from FCT and the company iCognitus4ALL - IT Solutions, Lda (grant number PD/BDE/127831/2016).

The authors acknowledge the participation of Mariana Pinote Moreira and Rita Vieira in the psychological assessments. Finally, we would also like to acknowledge the participants for their contribution to the study.

Authors' Contributions

CD contributed to conceptualization, data curation, formal analysis, investigation, methodology, and writing (original draft, review, and editing). PSC contributed to formal analysis, methodology, and writing (review and editing). NCS contributed to funding acquisition, supervision, and writing (review and editing). JMP contributed to funding acquisition, supervision, and writing (review and editing). All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Translation process of the User Satisfaction Evaluation Questionnaire items.
[DOCX File , 17 KB - mhealth_v9i6e19245_app1.docx]

References

1. Turner-McGrievy G, Jake-Schoffman DE, Singletary C, Wright M, Crimarco A, Wirth MD, et al. Using Commercial Physical Activity Trackers for Health Promotion Research: Four Case Studies. Health Promot Pract 2019 May 04;20(3):381-389 [FREE Full text] [doi: 10.1177/1524839918769559] [Medline: 29618233]
2. Steinert A, Haesner M, Steinhagen-Thiessen E. Activity-tracking devices for older adults: comparison and preferences. Univ Access Inf Soc 2017 Apr 8;17(2):411-419. [doi: 10.1007/s10209-017-0539-7]
3. Schlomann A, Seifert A, Rietz C. Relevance of Activity Tracking With Mobile Devices in the Relationship Between Physical Activity Levels and Satisfaction With Physical Fitness in Older Adults: Representative Survey. JMIR Aging 2019 Mar 06;2(1):e12303 [FREE Full text] [doi: 10.2196/12303] [Medline: 31518263]
4. Melin J, Bonn SE, Pendrill L, Trolle Lagerros Y. A Questionnaire for Assessing User Satisfaction With Mobile Health Apps: Development Using Rasch Measurement Theory. JMIR mHealth uHealth 2020 May 26;8(5):e15909 [FREE Full text] [doi: 10.2196/15909] [Medline: 32452817]
5. El-Amrawy F, Nounou MI. Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial? Healthc Inform Res 2015 Oct;21(4):315-320 [FREE Full text] [doi: 10.4258/hir.2015.21.4.315] [Medline: 26618039]
6. Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. BMC Public Health 2017 Nov 15;17(1):880 [FREE Full text] [doi: 10.1186/s12889-017-4888-1] [Medline: 29141607]
7. Rupp MA, Michaelis JR, McConnell DS, Smither JA. The role of individual differences on perceptions of wearable fitness device trust, usability, and motivational impact. Appl Ergon 2018 Jul;70:77-87. [doi: 10.1016/j.apergo.2018.02.005] [Medline: 29866329]
8. Kononova A, Li L, Kamp K, Bowen M, Rikard R, Cotten S, et al. The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change. JMIR mHealth uHealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
9. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older Adults' Acceptance of Activity Trackers. J Appl Gerontol 2017 Feb 07;36(2):127-155 [FREE Full text] [doi: 10.1177/0733464815624151] [Medline: 26753803]
10. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability Study of Mainstream Wearable Fitness Devices: Feature Analysis and System Usability Scale Evaluation. JMIR mHealth uHealth 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: 10.2196/11066] [Medline: 30409767]
11. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO Standards for Usability, Usability Reports and Usability Measures. In: Proceedings, Part I, of the 18th International Conference on Human-Computer Interaction. Theory, Design, Development and Practice. Cham: Springer; 2016 Jul Presented at: HCI 2016: Human-Computer Interaction. Theory, Design, Development and Practice; July 17-22, 2016; Toronto, ON p. 268-278. [doi: 10.1007/978-3-319-39510-4_25]
12. He Z, Kim S, Gong D. The Influence of Consumer and Product Characteristics on Intention to Repurchase of Smart band Focus on Chinese Consumers. International Journal of Asia Digital Art and Design Association 2017;21(1):13-18.
13. Chao C. Factors Determining the Behavioral Intention to Use Mobile Learning: An Application and Extension of the UTAUT Model. Front Psychol 2019 Jul 16;10:1652 [FREE Full text] [doi: 10.3389/fpsyg.2019.01652] [Medline: 31379679]
14. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1-16. [doi: 10.1186/2194-0827-1-1]

15. Brooke J. SUS: A 'Quick and Dirty' Usability Scale. In: Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996:184-194.
16. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. JMIR mHealth uHealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
17. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, Cross-Cultural Adaptation, and Validation of the Malay Version of the System Usability Scale Questionnaire for the Assessment of Mobile Apps. JMIR Hum Factors 2018 May 14;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
18. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. International Journal of Human-Computer Interaction 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
19. Gao M, Kortum P, Oswald F. Psychometric Evaluation of the USE (Usefulness, Satisfaction, and Ease of use) Questionnaire for Reliability and Validity. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2018 Sep 27 Presented at: Human Factors and Ergonomics Society Annual Meeting; October 1-5, 2018; Philadelphia, PA p. 1414-1418. [doi: 10.1177/1541931218621322]
20. Martins AI, Rosa AF, Queirós A, Silva A, Rocha NP. European Portuguese Validation of the System Usability Scale (SUS). Procedia Computer Science 2015;67:293-300. [doi: 10.1016/j.procs.2015.09.273]
21. Dantas C, Jegundo A, Quintas J, Martins A, Queirós R, Rocha N. European portuguese validation of usefulness, satisfaction and ease of use questionnaire (USE). Recent Advances in Information Systems and Technologies. WorldCIST 2017 Mar;570:561-570. [doi: 10.1007/978-3-319-56538-5_57]
22. Chin J, Diehl V, Norman K. Development of an instrument measuring user satisfaction of the human-computer interface. In: CHI '88: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1988 May Presented at: SIGCHI Conference on Human Factors in Computing Systems; Washington, DC; May 1988 p. 213-218. [doi: 10.1145/57167.57203]
23. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis J. USEQ: A Short Questionnaire for Satisfaction Evaluation of Virtual Rehabilitation Systems. Sensors (Basel) 2017 Jul 07;17(7):1589 [FREE Full text] [doi: 10.3390/s17071589] [Medline: 28686174]
24. Seifert A, Schlomann A, Rietz C, Schelling HR. The use of mobile devices for physical activity tracking in older adults' everyday life. Digit Health 2017 Nov 09;3:2055207617740088-2055207617740012 [FREE Full text] [doi: 10.1177/2055207617740088] [Medline: 29942617]
25. Farina N, Lowry RG. Older adults' satisfaction of wearing consumer-level activity monitors. J Rehabil Assist Technol Eng 2017 Oct 31;4:2055668317733258-2055668317733256 [FREE Full text] [doi: 10.1177/2055668317733258] [Medline: 31186939]
26. Lenz AS, Gómez Soler I, Dell'Aquilla J, Uribe PM. Translation and Cross-Cultural Adaptation of Assessments for Use in Counseling Research. Measurement and Evaluation in Counseling and Development 2017 Oct 03;50(4):224-231. [doi: 10.1080/07481756.2017.1320947]
27. World Health Organization. Process of translation and adaptation of instruments. 2007. URL: https://www.who.int/substance_abuse/research_tools/translation/en/ [accessed 2019-12-16]
28. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin 1959;56(2):81-105. [doi: 10.1037/h0046016]
29. Revythi A, Tselios N. Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning. Educ Inf Technol 2019 Feb 1;24(4):2341-2355. [doi: 10.1007/s10639-019-09869-4]
30. Lewis JR. The System Usability Scale: Past, Present, and Future. International Journal of Human–Computer Interaction 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
31. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
32. Myrberg K, Hydén LC, Samuelsson C. The mini-mental state examination (MMSE) from a language perspective: an analysis of test interaction. Clin Linguist Phon 2020 Jul 02;34(7):652-670. [doi: 10.1080/02699206.2019.1687757] [Medline: 31739693]
33. Montag C, Panksepp J. Primary Emotional Systems and Personality: An Evolutionary Perspective. Front Psychol 2017 Apr 11;8:464 [FREE Full text] [doi: 10.3389/fpsyg.2017.00464] [Medline: 28443039]
34. Yesavage JA, Brink T, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: A preliminary report. Journal of Psychiatric Research 1982 Jan;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4]
35. Guerreiro M, Silva A, Botelho M, Leitão O, Castro-Caldas A, Garcia C. Adaptação à população portuguesa da tradução do Mini Mental State Examination (MMSE). Rev Port Neurol 1994;1(9):240-248.
36. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User Acceptance of Wrist-Worn Activity Trackers Among Community-Dwelling Older Adults: Mixed Method Study. JMIR mHealth uHealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]

37. Mičková E, Machová K, Daďová K, Svobodová I. Does Dog Ownership Affect Physical Activity, Sleep, and Self-Reported Health in Older Adults? Int J Environ Res Public Health 2019 Sep 11;16(18):3355 [FREE Full text] [doi: 10.3390/ijerph16183355] [Medline: 31514379]
38. Venkatesh V, Bala H. Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
39. Kim H. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013 Feb;38(1):52-54 [FREE Full text] [doi: 10.5395/rde.2013.38.1.52] [Medline: 23495371]
40. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods 1996 Mar;1(1):16-29. [doi: 10.1037/1082-989x.1.1.16]
41. Makowski D. The psycho Package: an Efficient and Publishing-Oriented Workflow for Psychological Science. JOSS 2018 Feb;3(22):470. [doi: 10.21105/joss.00470]
42. Dziuban C, Shirkey E. When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological Bulletin 1974;81(6):358-361. [doi: 10.1037/h0036316]
43. Howard MC. A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? International Journal of Human-Computer Interaction 2015 Sep 14;32(1):51-62. [doi: 10.1080/10447318.2015.1087664]
44. Kyriazos TA. Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. PSYCH 2018;09(08):2207-2230. [doi: 10.4236/psych.2018.98126]
45. Osborne J, Costello A. Sample size and subject to item ratio in principal components analysis. Practical Assessment, Research, and Evaluation 2004;9(11):1.
46. O'Rourke N, Hatcher L. Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. Cary, NC: SAS Institute; 2013.
47. Patil VH, Singh SN, Mishra S, Donovan T. Parallel Analysis Engine to Aid Determining Number of Factors to Retain [Computer software]. 2007. URL: https://smishra.ku.edu/parallelengine.htm [accessed 2021-06-22]
48. Hooper D, Coughlan J, Mullen M. Structural Equation Modelling: Guidelines for Determining Model Fit. Electronic Journal of Business Research Methods 2008;6(1):53-60.
49. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. Methods of Psychological Research 2003;8(2):23-74. [doi: 10.4135/9781526421036946001]
50. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
51. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods 1999 Sep;4(3):272-299. [doi: 10.1037/1082-989x.4.3.272]
52. Şimşek GG, Noyan F. McDonald's ωt, Cronbach's α, and Generalized θ for Composite Reliability of Common Factors Structures. Communications in Statistics - Simulation and Computation 2013 Oct;42(9):2008-2025. [doi: 10.1080/03610918.2012.689062]
53. Trizano-Hermosilla I, Alvarado JM. Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. Front Psychol 2016 May 26;7:769 [FREE Full text] [doi: 10.3389/fpsyg.2016.00769] [Medline: 27303333]
54. Osborne J, Costello A, Kellow T. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005;10(7):86-99. [doi: 10.4135/9781412995627.d8]
55. Gadermann A, Guhn M, Zumbo B. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. Practical Assessment, Research, and Evaluation 2012;17(3):1-13. [doi: 10.7275/n560-j767]
56. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334. [doi: 10.1007/bf02310555]
57. Ferketich S. Focus on psychometrics. Aspects of item analysis. Res Nurs Health 1991 Apr;14(2):165-168. [doi: 10.1002/nur.4770140211] [Medline: 2047538]
58. Cohen J. A power primer. Psychological Bulletin 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
59. Schober P, Boer C, Schwarte LA. Correlation Coefficients. Anesthesia & Analgesia 2018;126(5):1763-1768. [doi: 10.1213/ane.0000000000002864]
60. Norman G. Likert scales, levels of measurement and the “laws” of statistics. Adv in Health Sci Educ 2010 Feb 10;15(5):625-632. [doi: 10.1007/s10459-010-9222-y]
61. Schneider F. Measuring Subjective Movie Evaluation Criteria: Conceptual Foundation, Construction, and Validation of the SMEC Scales. Communication Methods and Measures 2017 Jan 11;11(1):49-75. [doi: 10.1080/19312458.2016.1271115]
62. di Palo MT. Rating satisfaction research: Is it poor, fair, good, very good, or excellent? Arthritis Care & Research 1997 Dec;10(6):422-430. [doi: 10.1002/art.1790100610]
63. Thielsch MT, Thielsch C. Depressive symptoms and web user experience. PeerJ 2018 Feb;6:e4439 [FREE Full text] [doi: 10.7717/peerj.4439] [Medline: 29507832]

64. Li Q, Luximon Y. Understanding Older Adults' Post-adoption Usage Behavior and Perceptions of Mobile Technology. International Journal of Design 2018;12(3):93-110.

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
GDS: Geriatric Depression Scale
GFI: goodness of fit index
KMO: Kaiser-Meyer-Olkin
mHealth: mobile health
MMSE: Mini-Mental State Examination
PCA: principal component analysis
PSSUQ: Post-Study System Usability Questionnaire Usefulness
QUIS: Questionnaire for User Interaction Satisfaction
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
SUS: System Usability Scale
TAM 3: technology acceptance model
TLI: Tucker-Lewis index
USE: Satisfaction and Ease of Use
USEQ: User Satisfaction Evaluation Questionnaire
WHO: World Health Organization

Edited by G Eysenbach; submitted 10.04.20; peer-reviewed by K Liu, A Nguyen; comments to author 31.07.20; revised version received 24.09.20; accepted 10.10.20; published 29.06.21.

Please cite as:
Domingos C, Costa PS, Santos NC, Pêgo JM
European Portuguese Version of the User Satisfaction Evaluation Questionnaire (USEQ): Transcultural Adaptation and Validation Study
JMIR Mhealth Uhealth 2021;9(6):e19245
URL: https://mhealth.jmir.org/2021/6/e19245
doi:10.2196/19245
PMID:34185018

Original Paper

Sleep Detection for Younger Adults, Healthy Older Adults, and Older Adults Living With Dementia Using Wrist Temperature and Actigraphy: Prototype Testing and Case Study Analysis

Jing Wei[1,2], BEng, MASc; Jennifer Boger[2,3], BSc, MASc, PhD

[1]School of Computing and Information System, The University of Melbourne, Melbourne, Australia

[2]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[3]Schlegel Research Chair in Technology for Independent Living, Research Institute for Aging, Waterloo, ON, Canada

Corresponding Author:
Jennifer Boger, BSc, MASc, PhD
Department of Systems Design Engineering
University of Waterloo
Engineering 5, 6th Floor
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567 ext 38328
Email: jboger@uwaterloo.ca

Abstract

Background: Sleep is essential for one's health and quality of life. Wearable technologies that use motion and temperature sensors have made it possible to self-monitor sleep. Although there is a growing body of research on sleep monitoring using wearable devices for healthy young-to-middle-aged adults, few studies have focused on older adults, including those living with dementia.

Objective: This study aims to investigate the impact of age and dementia on sleep detection through movement and wrist temperature.

Methods: A total of 10 younger adults, 10 healthy older adults, and 8 older adults living with dementia (OAWD) were recruited. Each participant wore a Mi Band 2 (accemetry-based sleep detection) and our custom-built wristband (actigraphy and wrist temperature) 24 hours a day for 2 weeks and was asked to keep a daily sleep journal. Sleep parameters detected by the Mi Band 2 were compared with sleep journals, and visual analysis of actigraphy and temperature data was performed.

Results: The absolute differences in sleep onset and offset between the sleep journals and Mi Band 2 were 39 (SD 51) minutes and 31 (SD 52) minutes for younger adults, 49 (SD 58) minutes and 33 (SD 58) minutes for older adults, and 253 (SD 104) minutes and 161 (SD 94) minutes for OAWD. The Mi Band 2 was unable to accurately detect sleep in 3 healthy older adults and all OAWDs. The average sleep and wake temperature difference of OAWD (1.26 °C, SD 0.82 °C) was significantly lower than that of healthy older adults (2.04 °C, SD 0.70 °C) and healthy younger adults (2.48 °C, SD 0.88 °C). Actigraphy data showed that older adults had more movement during sleep compared with younger adults and that this trend appears to increase for those with dementia.

Conclusions: The Mi Band 2 did not accurately detect sleep in older adults who had greater levels of nighttime movement. As more nighttime movement appears to be a phenomenon that increases in prevalence with age and even more so with dementia, further research needs to be conducted with a larger sample size and greater diversity of commercially available wearable devices to explore these trends more conclusively. All participants, including older adults and OAWD, had a distinct sleep and wake wrist temperature contrast, which suggests that wrist temperature could be leveraged to create more robust and broadly applicable sleep detection algorithms.

(JMIR Mhealth Uhealth 2021;9(6):e26462) doi:10.2196/26462

KEYWORDS

sleep monitoring; wearables; accelerometer; wrist temperature; circadian rhythm; younger adults; older adults; dementia; mobile phone

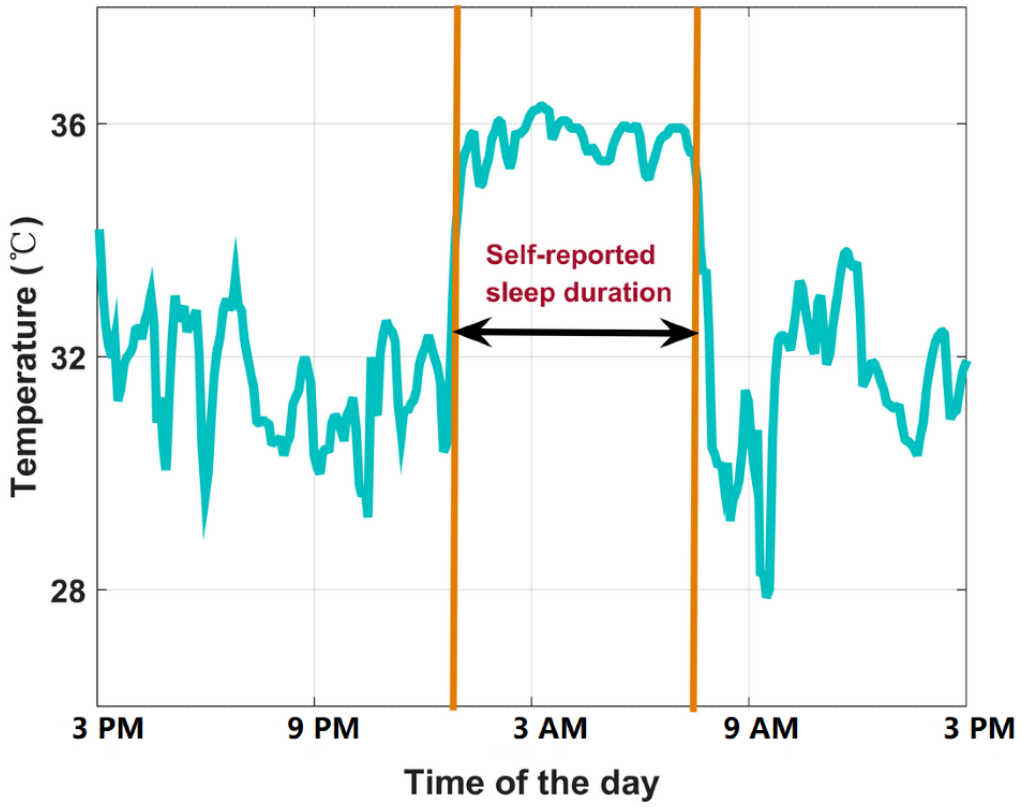
Introduction

On average, we spend one-third of our lives asleep [1]. Quality and quantity of sleep regulate daytime behaviors and functions as well as significantly impact health and well-being. Sleep has been shown to affect day-to-day memory and concentration [2-4]. Lifetime sleep habits appear to correlate with one's likelihood of having conditions, such as Alzheimer disease [5-7] and cardiovascular diseases [8,9]. Although the benefits of getting a good night's rest have received extensive research, there is still much we do not know about the differences between how people sleep and what can be done to support better sleep.

Understanding sleep patterns is a complex undertaking, as sleep is impacted by conscious and subconscious control by the individual as well as environmental factors. Standard clinical sleep evaluation typically uses polysomnography (PSG), which is considered the *gold standard* in sleep studies [10]. PSG requires the person being assessed to sleep in a laboratory with devices attached to the body in single or infrequent sessions. Although the data collected by PSG can be used to diagnose many medical conditions, its intrusive, unnatural nature might not reflect people's usual sleep quality and is not appropriate for long-term or frequent monitoring of sleep [11].

Actigraphy (ie, accelerometer sensors and gyroscopes) is being increasingly used to measure people's sleep and activities by estimating related parameters, such as sleep onset and offset [12,13]. Many clinical studies have adopted actigraphy to measure circadian rhythm, which is a major factor in regulating people's sleep and wake rhythms [14,15]. Driven by circadian rhythms, many biological processes, including core body temperature (CBT), have 24-hour diurnal variations. During sleep, people's CBT drops about 1 °C [16]. Thermal regulation inside the human body causes wrist temperature to exhibit an opposite pattern to CBT [17]; wrist temperature increases before people fall asleep, remains at a relatively high level when people are asleep, and then drops in the morning when people wake up (Figure 1). CBT has been used extensively to study circadian rhythms; however, it usually requires invasive gut or rectal temperature measurements [18]. As an alternative, wrist temperature can be measured in daily life and is found to be more correlated with sleeping status than CBT [19]. Recent studies have provided increasing evidence that wrist temperature can be used for sleep monitoring [17,20-22].

Figure 1. Example wrist temperature over a 24-hour period for a healthy sleep pattern.



Low sleep quality and irregular sleep patterns become more prevalent with increasing age and dementia [23-26]. For example, poor nighttime sleep has been shown to be a key factor in older adults' tendency to have more daytime sleepiness [27], which has been associated with changes in circadian rhythms in older adults [28]. Sleep can also be used as an indicator of

comorbidities, such as sleep apnea and restless leg syndrome, both of which become more prevalent with increasing age [29,30]. Disrupted sleep patterns have been found to be even more prevalent in people with dementia [31]. Poor sleep has been found to be predictive of more severely impacted cognition in older adults, including people with dementia [32] and older women without dementia [33]. Therefore, the ability to monitor older adults' sleep and circadian rhythms on an ongoing basis is increasingly being used to understand and support their sleep, which, in turn, supports health and well-being [34].

The proliferation of smart wearable technologies has contributed to the substantial growth in technology adoption among older adults; older adults are the most rapidly increasing group of users of new technologies [35,36]. One of the most commonly adopted technologies is smart wearable devices, which usually have built-in accelerometer and gyroscope sensors to detect people's sleep (eg, Fitbit, Garmin, and Mi Band) [37-39]. Commercial smart wristbands have been designed to output several sleep parameters, including sleep onset, offset, duration, and wake after sleep onset, and some wristbands also give a sleep quality score. Sleep parameters can be synchronized to one's smartphone via Bluetooth (and then to the cloud, should they wish) so that users can easily access their sleep data. Relative affordability and unobtrusiveness contribute to the growing popularity of commercial smart wristbands. In addition, smart wearable devices are being increasingly and widely used in health research [40-42]. For example, Gibson et al [43] evaluated the reliability of actigraphy to measure the sleep of people living with dementia.

Although most smart wristbands can provide sleep monitoring and analysis, they fail to monitor people's sleep from an internal perspective; they cannot measure circadian rhythm directly. As discussed earlier, wrist temperature is a reflection of CBT, which is indicative of a person's biological circadian rhythm. As such, wrist temperature could have value as an addition to accemetry-based wearable sleep monitoring systems, which could then be used to support better understanding and management of sleep. To be successful, these systems should work for different stakeholder groups, including older adults and older adults living with dementia (OAWD). Therefore, this research was guided by the question, "How do sleep and wrist temperature patterns compare for younger adults and older adults, including OAWD?"

Methods

Participants

This study measured sleep patterns and wrist temperature patterns of 3 groups of participants: (1) healthy younger adults (aged 20-30 years), (2) healthy older adults (aged≥65 years), and (3) OAWD. Inclusion criteria were any gender, any race, age in the targeted populations, and ability to understand English. An additional criterion for OAWD was early- to middle-stage dementia. Individuals diagnosed with severe sleep disorders were excluded from the study.

We recruited 10 participants for each group based on the convenience sampling method. Healthy adults were recruited through university-wide poster advertisements and word of mouth. Healthy older adults were recruited through phone calls and email invitations. OAWD participants were recruited from a local long-term care (LTC) residence. The staff at the LTC setting requested that they coordinate the initial identification of potential participants and support the recruitment process. Each potential participant was then contacted and introduced to the study by the researcher. Assent to continue participating in the study was obtained from each OAWD participant every time they were contacted by the researchers. This study was approved by the Human Research Ethics Committee of the University of Waterloo (ORE #31860 and ORE #40459).

Equipment

Each participant was asked to wear a commercially available smart wristband (Mi Band 2, XiaoMi; Figure 2) and a custom-built wristband developed by the authors (Figure 3) on their left arm for 14 days. The Mi Band 2 was chosen because of its popularity, low cost, long battery life, and accessibility of actigraphy data. The custom-built band consisted of a temperature sensor (iButton DS1922L, Maxim) that maintained contact with the anterior side of the wrist and a 3-axis accelerometer sensor (AX3, Axivity). Both sensors operated offline by storing data locally; data were extracted after the 2-week study period using USB adaptors. All participants were offered to keep the Mi Band 2 if they wished on completion of the study.

Figure 2. The Mi Band 2.

Figure 3. The customized wristband.

Study Protocol

The study protocol was the same for healthy younger adults and healthy older adults, with an altered OAWD protocol. Before the start of the study, all healthy participants were asked to complete 4 questionnaires: (1) a demographic form; (2) morningness-eveningness questionnaire, which is a measure of *circadian rhythm type* [44]; (3) Pittsburgh Sleep Quality Index (PSQI), which is a measure of subjective sleep quality [45]; and (4) Epworth Sleepiness Scale (ESS), which is a measure of daytime sleepiness [46]. Participants were then given the 2 wristbands; instructed on how to wear them; and asked to wear both on their left wrist for 14 days, except when showering or bathing. Participants were also asked to fill out a daily sleep journal that was adapted from a study by Monk et al [47] to capture sleep onset, offset, and subjective sleep quality.

As OAWD were not able to fill out the questionnaires, only the demographic form was completed and was done with the help of staff from their LTC residence. In addition, the Mini-Mental State Examination (MMSE) [48] was conducted to approximate the level of cognition for each OAWD. After giving each OAWD the wristbands, their personal support workers (PSWs) were asked to help with getting wristbands off and on when they assisted OAWD shower or bathe as well as to routinely check that the person was not adversely affected by wearing the wristbands. In addition, PSWs were asked to observe participants' sleep status and check off either *awake* or *sleeping* on a sheet every half an hour throughout the day and night as part of their routine in with the OAWD. Sleep status from this sheet was then used to generate an observed sleep journal.

Ground Truth of Sleep Parameters

Sleep onset and offset, the number of nighttime wake-ups, and sleep quality scores were obtained from the sleep journals and were used to help interpret actigraphy and temperature data. Although we thought it was unlikely that people could estimate

their sleep at a minute level, we considered sleep journals of younger adults and healthy older adults reliable at the half-hour interval level. Occasionally, participants would miss reporting sleep onset and offset. In these cases, the sleep onset and offset were manually identified from the data from the custom-built wristband AX3 actigraphy data.

For OAWD, the intention was to use the sleep status sheet filled out by PSWs to extract sleep parameters and use these as ground truths for sleep. However, at the end of the study, it was found that, on average, 40% of the sheet was filled (ie, the PSWs did not report sleep status of OAWD 60% of the time). Owing to the large amount of data missing from the OAWD sleep journals, we chose to manually identify sleep periods from the AX3 data and combined those identified sleep periods with PSW sheets to estimate sleep for all OAWD.

Feature Extractions

Sleep Parameter Extraction

Sleep onset, sleep offset, and sleep duration from the sleep journals were compared with the Mi Band 2 for daily sleep monitoring; the mean values across the 14 days were calculated and are presented in Table 1. The absolute differences between the Mi Band 2 and journal parameters were used to indicate the agreement of sleep detection between Mi Band 2 and self-reported sleep.

Table 1. Sleep parameters calculated from sleep journals (younger and older adults) and recreated sleep journals (older adults living with dementia).

Sleep parameters	Younger adults	Older adults	Older adults living with dementia
Onset, mean (SD)	12:32 AM (53 min)	11:49 PM (61 min)	9:37 PM (105 min)
Offset, mean (SD)	7:55 AM (54 min)	7:31 AM (78 min)	7:05 AM (40 min)
Duration (min), mean (SD)	442 (26)	462 (88)	562 (123)

Sleep Periods Identification of OAWD

As people generally have minimal movements during sleep, sleep periods can be manually identified by segments of raw AX3 data that have minimal variations. For one night's sleep, multiple sleep periods could be identified if the person had multiple wake-ups. Once sleep periods were estimated, the sleep onset, offset, and duration were calculated for the longest least varied data period. If participants woke up during the night and remained awake for a long period (ie, more than 2 hours), this period was not counted as sleep duration. AX3-estimated sleep was combined with the partially completed observational sheet by PSWs to create a recreated sleep journal for OAWD.

Wrist Temperature Feature Extraction

As participants took the customized wristband off to shower, all data points lower than 28 °C were removed and interpolated using the nearest values. The temperature data were then smoothed by applying a median filter with a window size of 3. Descriptive statistics were calculated to provide a profile of the general characteristics of the participants' wrist temperature patterns and rhythm indices related to circadian rhythms. The mean and SD were calculated for the wrist temperature rhythm of every 24-hour period that began at 3 PM because this segmentation included full nighttime sleep periods and daytime naps. On the basis of sleep onset and sleep offsets obtained from sleep journals, the mean temperatures were calculated for sleep time and wake time. The mean wake time temperature was subtracted from the mean sleep time temperature to calculate the mean sleep and wake temperature difference, which was used as a measure of the extent of temperature changes between the sleep and wake states in this study.

Interday stability (IS) and intraday variability (IV) were calculated to examine the regularity and rhythmicity of wrist temperature rhythms for each participant [49]. IS has been used to measure the stability of rhythms across several consecutive days, whereas IV reflects the fragmentation of the wrist temperature rhythm of every 24-hour period. IS and IV were calculated using Equations (1) and (2), respectively:

where N is the total number of wrist temperature data points, p is the number of wrist temperature data points per day (in this study, P=288 as the iButton is sampled every 5 min), X is the mean of all temperature data, X_h is the hourly mean of wrist temperature, and X_i is the data point at time i. IS and IV were calculated for each participant for the total study period.

Statistical Analysis

One-tailed t test was used to compare the sleep parameters and rhythm indices between different groups (young vs old and old vs OAWD). The Pearson correlation coefficient was used to assess correlations between different circadian rhythms and sleep parameters calculated; P<.05 was considered to be significant.

Case Study Analysis

When analyzing the collected data, it was found that the sleep of some participants was not correctly reported by the Mi Band 2. For example, most PSWs reported that OAWD had 8 hours of sleep at night, but the Mi Band 2 detected as few as 40 minutes of sleep. The AX3 data from the custom-built band were inspected for the periods Mi Band 2 had classified as *wake periods* during the night as well as wrist temperature measurements for the same time. A healthy younger participant who had good sleep quality; a healthy older participant with poor sleep quality; and 2 OAWD who had poor, interrupted sleep were selected as case studies to illustrate this phenomenon and are presented later in this paper.

Results

Demographics and Sleep Patterns

In total, we recruited 10 healthy younger adults, 10 healthy older adults, and 8 OAWD. A total of 7 OAWD were able to give their own consent and sign the consent form; for the 1 participant who could not self-consent, his power of attorney signed the consent form, and assent was obtained from the participant. Participants' demographic data are presented in Table 2. The average MMSE score of the 8 OAWD participants was 20. Moreover, 7 OAWD were in the mild cognitive impairment category (ie, a score between 18 and 23), and 1 OAWD had a score of 16, which indicates severe cognitive impairment [50].

Table 2. Participants' demographics.

Demographic	Healthy younger adult group (n=10)	Healthy older adult group (n=10)	Older adults living with dementia group (n=8)
Age (years), mean (SD)	24.1 (2.23)	75 (7.36)	83.25 (10.38)
Sex, n (%)			
Male	4 (40)	4 (40)	4 (50)
Female	6 (60)	6 (60)	4 (50)
BMI (kg/m^2), mean (SD)	23.69 (6.30)	25.94 (3.11)	29.90 (8.91)
Pittsburgh Sleep Quality Index, mean (SD)[a]	5.2 (2.39)	7.8 (3.56)	—[b]
Mini-Mental State Examination, mean (SD)[c]	—	—	20 (2.39)

[a]The Pittsburgh Sleep Quality Index was not measured for the older adults living with dementia group as they cannot reliably complete this measure.

[b]Test was not administered as the test was not appropriate for the participant group.

[c]The Mini-Mental State Examination of the healthy younger adult group and the healthy older adult group as they were cognitively intact.

The sleep parameters obtained from the sleep journals for all participants are presented in Table 1. Compared with older adult groups, healthy younger participants tended to sleep later and wake up later, with the shortest average sleep duration among the 3 groups. Between the 2 older adult groups, OAWD slept earlier and had a longer sleep duration. Although we did not require participants to record daytime naps, we asked them to report whether they had the habit of napping. Only 1 younger adult and 1 healthy older adult reported that they took regular noontime naps. All OAWD napped during the day, sometimes more than once. As the recreated sleep journals of OAWD only summarized nighttime sleep and did not capture daytime naps, the total sleep duration of OAWD might be longer than reported. Healthy older adults had better compliance in reporting and provided more detailed sleep journals (eg, a few participants consistently recorded their nap time during the day).

Comparison Between Mi Band 2 and Sleep Journal

The absolute differences in sleep parameters between the Mi Band 2 and sleep journals are summarized in Table 3, and significant differences were found in sleep duration between the younger adult group and the healthy older adult group. Between the 2 older adult groups, the absolute differences in sleep onset, offset, and duration of OAWD were significantly higher than those in the healthy older adult group. As shown in Table 4, of the 106 days of data measured from the 8 OAWD, for 28 of the days (28/106, 26% of days monitored), the Mi Band 2 did not detect any sleep, and we confirmed that all OAWDs were wearing the Mi Band 2 correctly throughout the study. Discrepancies between sleep parameters reported in the recreated sleep journal and detected by Mi Band 2 for OAWD were large, namely, sleep onsets detected by Mi Band 2 were often noticeably delayed (ie, on the order of hours), and the offset was earlier than what was reported. The average absolute difference in sleep duration between the Mi Band 2 data and the recreated sleep journal for OAWD was more than 6 hours. A discussion of this detection error is presented later in the *Case Studies* section.

Table 3. The absolute difference of sleep parameters between (recreated) sleep journals and the Mi Band 2 for each group.

Sleep parameters	Healthy younger adults	Healthy older adults	Older adults living with dementia
Onset			
Difference (min), mean (SD)	39 (51)	49 (58)	253 (104)
P value	—[a]	—	<.001
Offset			
Difference (min), mean (SD)	31 (52)	33 (58)	161 (94)
P value	—	—	<.001
Duration			
Difference (min), mean (SD)	49 (59)	64 (77)	379 (163)
P value	—	.04	<.001

[a]No significance.

Table 4. Number of days with sleep reported by the Mi Band 2 for older adults living with dementia.

ID	Total days (n)	Valid days, n (%)
OAWD[a] 1	14	14 (100)
OAWD 2	13	2 (15)
OAWD 3	10	10 (100)
OAWD 4	14	7 (50)
OAWD 5	14	12 (86)
OAWD 6	13	8 (62)
OAWD 7	14	11 (79)
OAWD 8	14	14 (100)
Total	106	78 (74)

[a]OAWD: older adults living with dementia.

Wrist Temperature Rhythm Comparison

The average wrist temperature data of the 3 groups are shown in Figure 4. Although all curves have higher daytime (awake) temperatures, the OAWD group has the flattest temperature curve with the lowest nighttime wrist temperature and the highest daytime wrist temperature compared with the other groups. Although the average nighttime wrist temperature was similar for healthy older and younger adults, older adults tended to have a higher daytime wrist temperature, causing a flatter average curve for healthy older adults compared with younger adults. The wrist temperature indices are presented in Table 5.

Figure 4. Average wrist temperature patterns for healthy younger adults, healthy older adults, and older adults living with dementia. Plots are expressed as the mean temperature (standard errors of the mean). OAWD: older adults living with dementia.

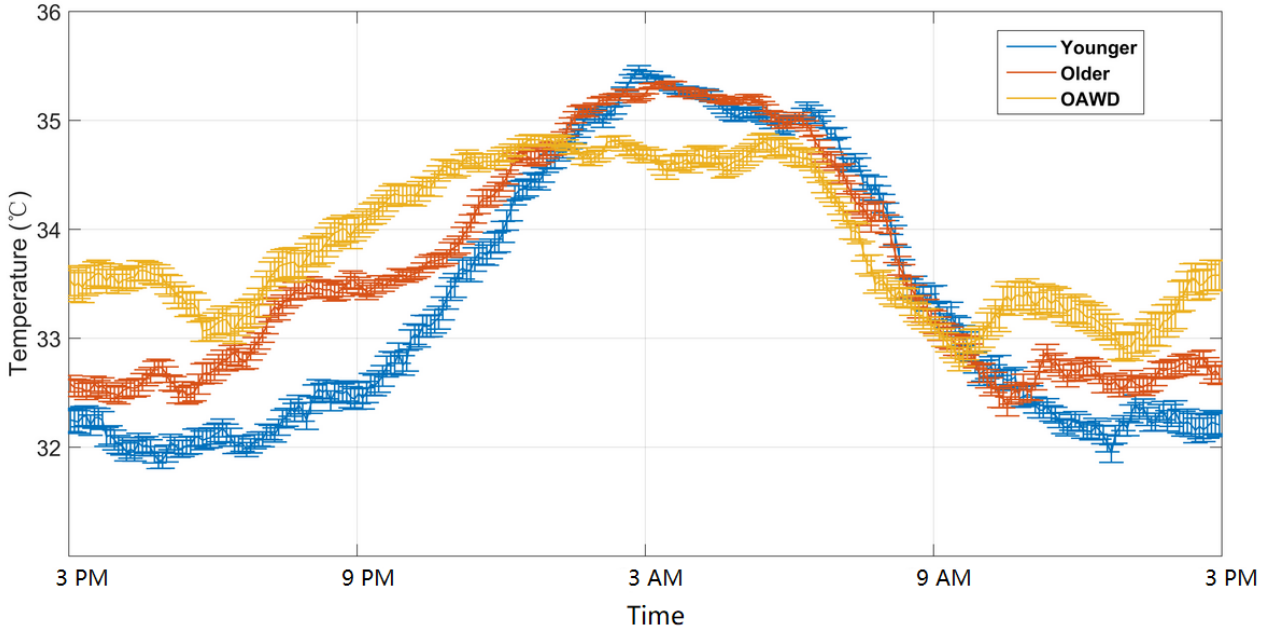


Table 5. Wrist temperature indices.

Parameters	Healthy younger adults (n=10)	Healthy older adults (n=10)	Older adults living with dementia (n=8)
Average wrist temperature parameters (°C), mean (SD)			
Sleep temperature	34.98 (0.55)	34.97 (0.34)	34.66 (0.53)
Wake temperature	32.49 (0.83)	32.94 (0.67)	33.40 (1.03)
Sleep and wake temperature difference	2.48 (0.88)	2.04 (0.70)	1.26 (0.82; *P*=.02)
Cosinor analysis			
MESOR[a] (°C), mean (SD)	33.34 (0.61)	33.65 (0.46)	33.90 (0.84)
Amplitude (°C), mean (SD)	1.72 (0.57)	1.45 (0.50)	0.93 (0.59; *P*=.03)
Acrophase (min), mean (SD)	246 (59)	183 (91; *P*=.04)	44 (145; *P*<.001)
Nonparametric analyses (°C), mean (SD)			
Interday stability	0.53 (0.10)	0.52 (0.16)	0.32 (0. 19; *P*=.02)

[a]MESOR: midline-estimating statistic of rhythm.

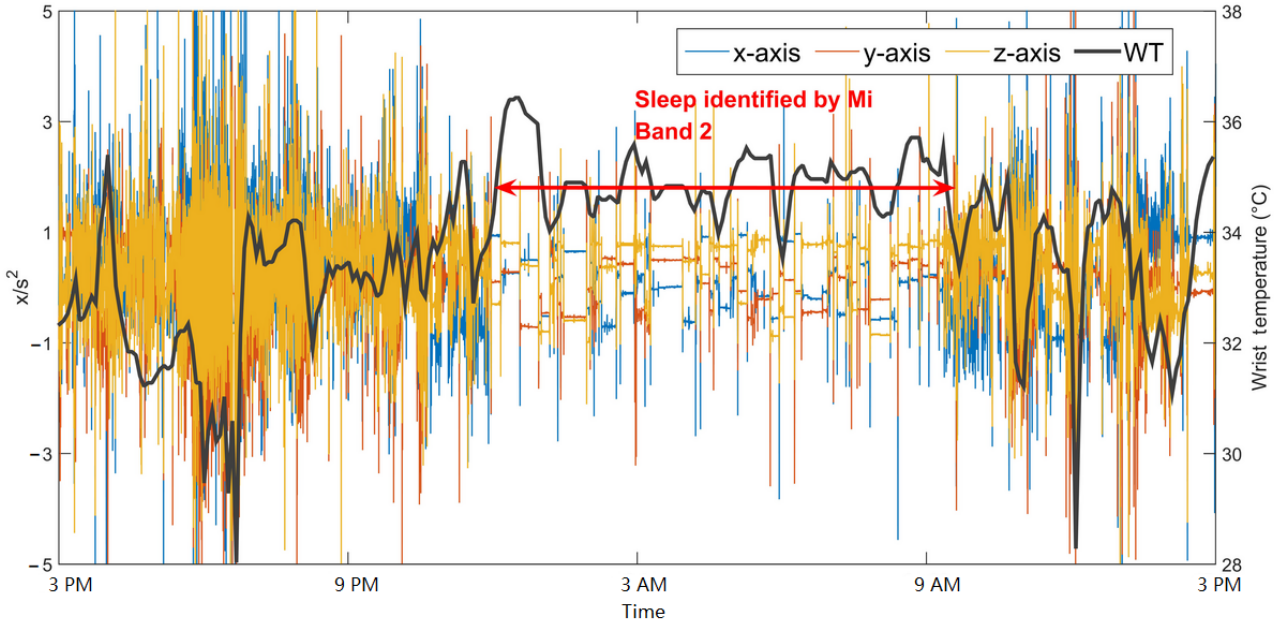
Case Studies

Data from 4 participants are presented below as illustrative case studies. One of the case studies is of a younger *healthy sleeper*, and the other 3 cases are representative data from *irregular sleepers*, 1 of which is a healthy older adult and the other 2 are OAWD.

Case I: Healthy Younger Adult With Regular Sleep

The healthy younger adult (YA1) was a 25-year-old man (PSQI=5 and ESS=3) who did not report any sleep disorders. One day of YA1 data, including wrist temperature, AX3 data, and sleep onset or offset detected by Mi Band 2, is shown in Figure 5. The sleep period and wake period can be distinguished using AX3 data, which shows that YA1 was mostly static with a relatively high and stable wrist temperature during sleep. The sleep onset and offset detected by the Mi Band 2 align well with changes in AX3 and wrist temperature data.

Figure 5. Example wrist temperature and AX3 data for a healthy younger adult with good sleep (YA1). Sleep onset and offset detected by Mi Band 2 are indicated by the double red arrows. WT: wrist temperature.

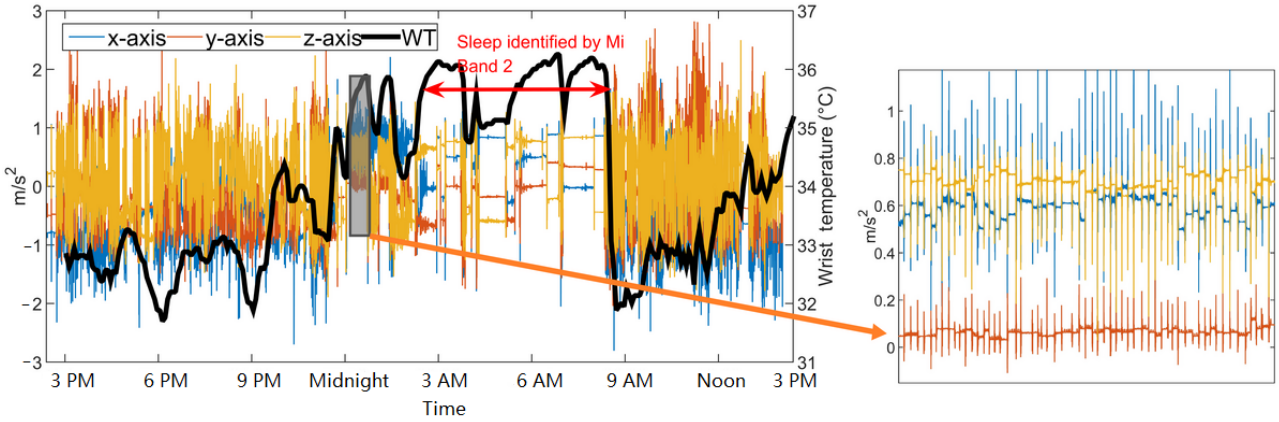


Case II: Healthy Older Adult With Irregular Sleep

The healthy older adult (OA1) was an 83-year-old man (PSQI=6 and ESS=7) who did not report any sleep disorders. OA1 was considered to be an *irregular sleeper* based on his data; one day of representative data is shown in Figure 6. Sleep onsets detected by the Mi Band 2 were, on average, 2 hours later than OA1's self-reported sleep onset.

For the day represented in Figure 6, OA1 self-reported falling asleep at around 11:45 PM, whereas the Mi Band 2 reported that his sleep started at 2:43 AM. Significant movements were captured by AX3 between 12:00 AM and 2:30 AM; a magnified image, except for these data, is shown on the right side of Figure 6, where multiple peaks can be observed in all 3 axes. After 2:30 AM, the AX3 data became more static and similar to the AX3 data collected from *regular sleepers*. The movements between 12:00 AM and 2:30 AM were very periodic and occurred approximately every 30 seconds for most of the 2.5-hour time frame; during this time, the wrist temperature remained relatively high, and the participant reported being asleep. As it seems unlikely that these movements were made by the participant consciously and people cannot *fake* body temperature changes, the participant was likely asleep but experiencing irregular body movements.

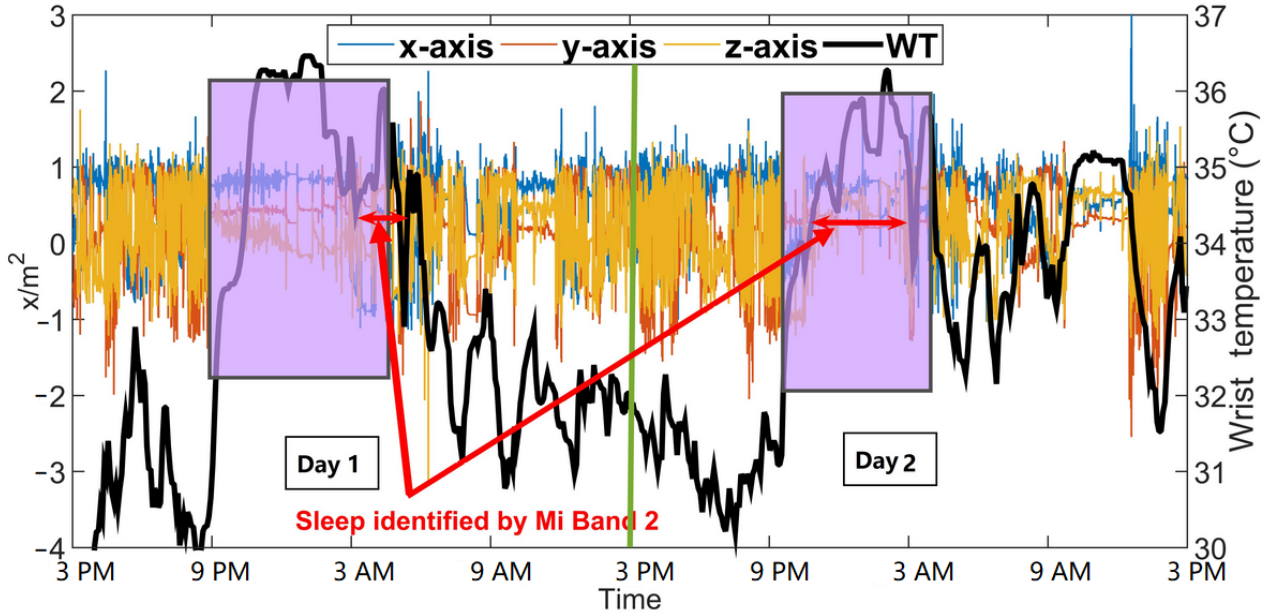
Figure 6. Example wrist temperature and AX3 data for a healthy older adult with irregular sleep (OA1). Sleep detected by the Mi Band 2 is indicated by the red double arrow. WT: wrist temperature.



Case III: OAWD1 With Parkinson Disease and Sleep Apnea

OAWD1 was a 68-year-old man (MMSE score=21, diagnosed with Parkinson disease and sleep apnea); 2 days of data are shown in Figure 7. Compared with OA1 (Figure 6), OAWD1's data (both wrist temperature and AX3 data) have a more evident contrast between day and night. Even so, only parts of his sleep were detected by Mi Band 2. As shown in the left portion of Figure 7, there was little movement between 9:00 PM and 3:00 AM, and the wrist temperature was relatively high, but this time frame was not identified as sleep by Mi Band 2.

Figure 7. Two consecutive days of wrist temperature and AX3 data for older adults living with dementia 1. Sleep onset or offset were obtained from the recreated sleep journal. Sleep detected by the Mi Band 2 is indicated by the red double arrow. The sleep of 2 nights in the recreated sleep journal is labeled by the shaded boxes. WT: wrist temperature.



Case IV: OAWD2 With Sleep Apnea and Insomnia

OAWD2 was a 68-year-old woman (MMSE score=19) who was diagnosed with sleep apnea and insomnia and was an *irregular sleeper*. Before the start of the study, OAWD2's PSW reported that she had very poor sleep (ie, would wake up at night frequently and sometimes cannot fall asleep at all). During her 13 days in our study, the Mi Band 2 indicated no sleep for 11 days; PSWs confirmed that she wore the wristbands the entire time. Figure 8 shows 2 consecutive days of wrist temperature and AX3 data. The left portion of Figure 8 shows data where no sleep was detected by Mi Band 2, and the right portion shows a short period of sleep detected by Mi Band 2. Compared with data from OA1 in Figure 6, there are no obvious *static periods*, and the day and night contrast for movement is relatively indistinguishable. Although there are differences in daytime and nighttime wrist temperatures, these are more difficult to distinguish than in other participants.

A close-up of an excerpt of AX3 data for day 1 is shown in Figure 9, where 8 short episodes of static data are highlighted. Each highlighted episode lasted <30 minutes, and considerable movement was observed between episodes. These data indicate that the participant may have had some sleep, but if so, the sleep was in short, fragmented periods. As the participant was diagnosed with sleep apnea, this could partially account for frequent nighttime wake-ups.

Figure 8. Two consecutive days of wrist temperature and AX3 data of older adults living with dementia 2. Sleep onset or offset were obtained from the recreated sleep journal. WT: wrist temperature.

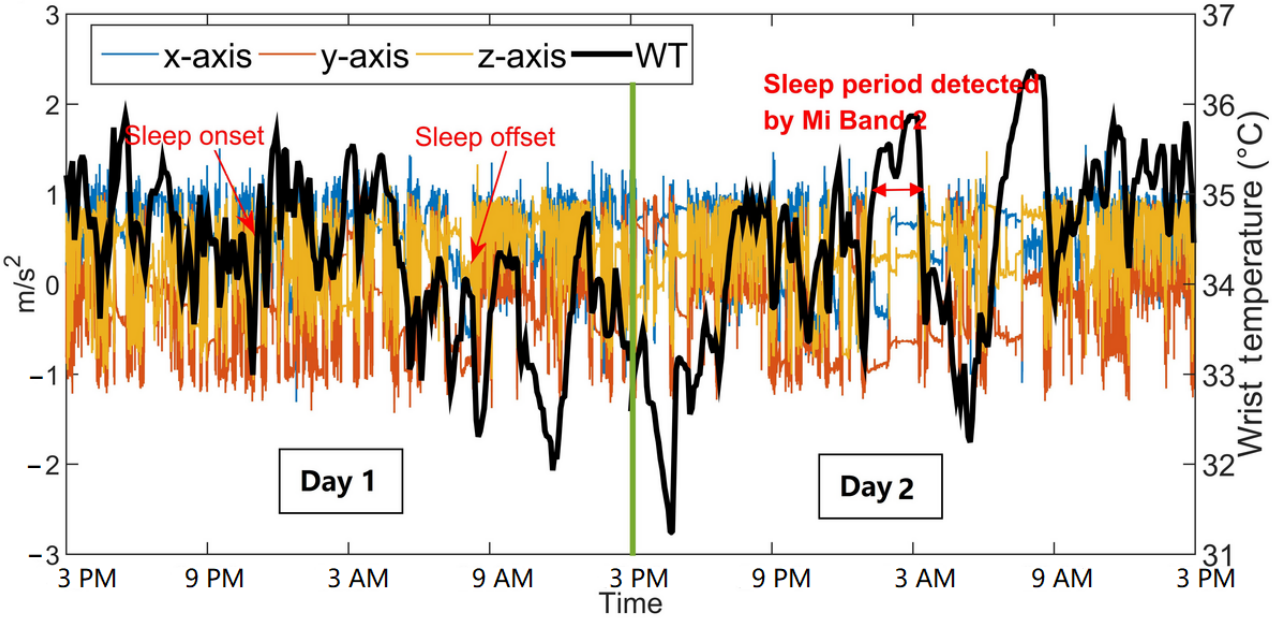
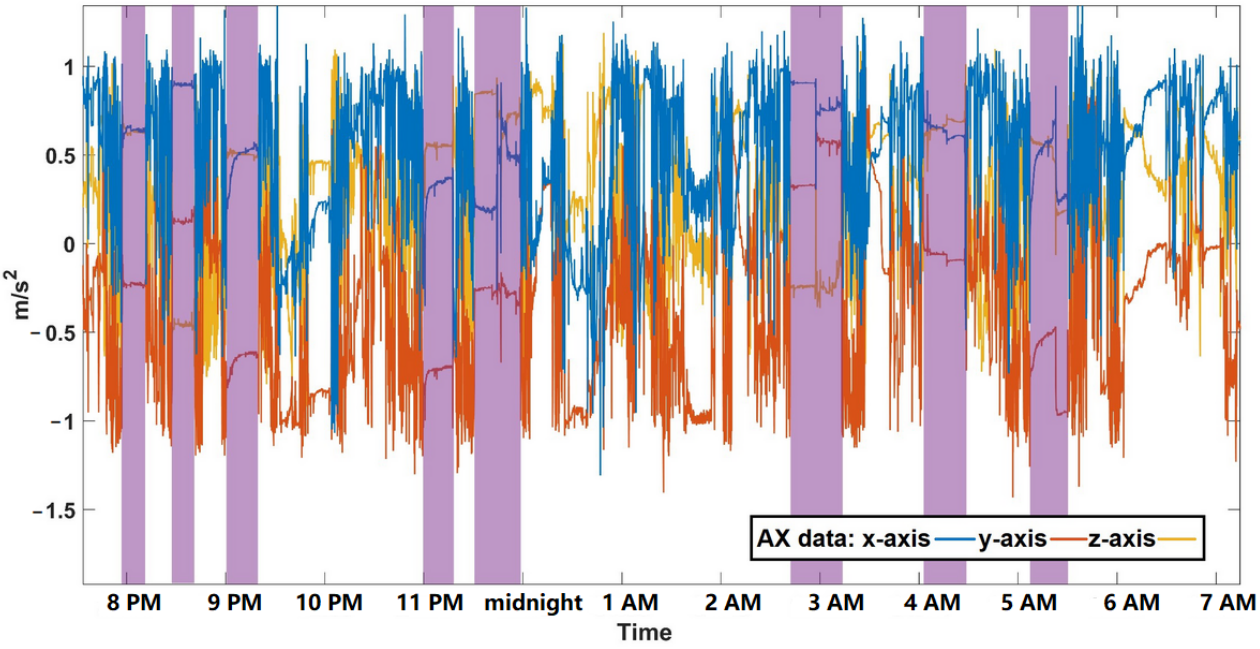


Figure 9. An 11-hour period (between 8 PM and 7 AM) of AX3 data for older adults living with dementia 2. Periods of relatively static data are shown by purple rectangles. WT: wrist temperature.



Discussion

Principal Findings

In line with other recent studies [17,19], this study found that wrist temperature increased when people were asleep, even when their sleep patterns were irregular. As illustrated in the case studies presented earlier, wrist temperature appears to correspond with sleep status changes regardless of whether body movements occurred; wrist temperature increased when people were asleep and decreased when they woke up. This adds evidence that wrist temperature is correlated with circadian rhythms and associated changes in thermal regulation of the body during sleep. This appears to hold true for OAWD as well, whose circadian rhythms are impacted by both aging and dementia (and often other morbidities); their wrist temperature appears to reflect when they are asleep and awake. A larger sample size is needed to draw more definitive conclusions regarding the impact dementia may have on CBT and wrist temperature.

From the results presented in Table 5 and Figure 4, there seems to be a trend of increasing wake temperature and decreasing sleep and wake temperature difference as people age and people become more pronounced dementia in this study. It is not clear from our data whether and how much this increased discrepancy is due to dementia or other factors, such as increased average age of OAWD, greater prevalence of comorbidities, and/or how living in an (often busy) LTC environment may disrupt sleep. The less significant day versus night wrist temperature contrast suggests changes in the circadian rhythms of older adults and OAWD compared with younger adults. This finding is consistent with other studies (eg, [51]).

The IS values suggest that the 2 healthy groups had similar IS for their wrist temperature rhythm, and the dementia group had lower wrist temperature rhythm stability. The lower IV values of the older adult groups indicate that changes in wrist temperature rhythms tend to be less variable intraday as people get older, especially with people who have dementia. As IS reflects the stability and the IV reflects the fragmentation of one's wrist temperature rhythm, the significantly lower IS and IV of OAWD could be associated with irregular sleep patterns and increased daytime sleepiness. Most OAWD in this study took frequent naps and experienced insomnia at night, which almost certainly increased the average daytime wrist temperature, lowered the average nighttime wrist temperature, and accounted for a greater amount of flux in nighttime wrist temperature. Rhythm indices, such as IS and IV, could be further explored with these and other populations to determine how they might be used to infer circadian rhythm health.

Some studies [30,52-54] have reported excessive body or wrist movements of OAWD during sleep. This research supports the hypothesis that advanced sleep monitoring systems should detect nighttime body movements. Although the exact cause of abnormal body movements by the older adults in this study is unknown, body movement during sleep seems to become increasingly common with age and was present in all our OAWD participants. A system that can detect involuntary movements could be valuable in the detection of previously unknown conditions (such as the case with our OA1 participant) and possibly support the diagnosis and management of conditions that impact sleep.

There are 2 noteworthy findings regarding sleep detection using the Mi Band 2. First, by comparing the results of the Mi Band 2 with AX3 data from the custom-built wristband and reported sleep from sleep journals, it appears that the Mi Band 2 does not accurately detect people's sleep when they have pronounced or irregular body movements during sleep. Second, the Mi Band 2 sleep detection algorithm does not appear to work well if the user has fragmented sleep (ie, waking up multiple times during the night). For example, OAWD2 slept in short and sporadic

episodes; most of her sleep was undetected by Mi Band 2, which reported sleep for only 2 of her 13 recorded days. From the example data in Figure 9, we can observe that each sleep episode of OAWD2 was short and each wake-up accompanied body movements and that the Mi Band 2 did not detect any sleep. As the Mi Band 2 detects sleep based on accemetry data, it is plausible that periodic body movements during sleep, such as those seen in the older adults presented in the case studies, may be the cause of the misclassification of these periods as *awake* instead of *asleep*. Moreover, the Mi Band 2 may be trained on data from people with 1 or 2 long nighttime sleeping periods than multiple short ones. The hypothesis that movements and short sleeping episodes cause the Mi Band to miss episodes of sleep is supported by the data from our other participants.

Although the sleep detection algorithms for most commercial smart wristbands are proprietary information, as they rely on accelerometers and gyroscopes, their sleep detection is based on movement. For a regular sleeper, frequent and substantial body movements at night usually indicate the person is awake, and fewer and infrequent movements indicate sleep. Although an activity-based rule for sleep detection may be robust for people who are mostly still when they sleep, this study shows that sleep detection based on activity is not accurate for some older adults, especially OAWD. As younger adults are the primary consumers of wearable wristbands, the Mi Band 2 sleep detection algorithms may have been trained predominantly on data from younger adults with healthy sleep patterns, thus biasing the algorithm to detect sleep for people who fit that profile. This aligns with other studies that have found that commercial wristbands performed poorly on populations with sleep disorders [55]. Therefore, other smart wristbands that track sleep based on movement alone may have similar performance as the Mi Band 2 for the same reasons; however, this was not investigated in this study; therefore, it remains a question for future research.

This research highlights the need for the older adult population to be included in the development of sleep detection algorithms, including the possibility of developing smarter algorithms that autonomously identify the type of sleeper a person is and self-adjust appropriately to that person's sleep habits. This requires research into how to categorize what *good* sleep is for that particular person in a way that helps people, their care partners, and/or caregivers understand. This information could then be used to better manage sleep without missing information that may indicate a change in sleep or sleep patterns that are of possible concern. As wrist temperature appears to be highly correlated with sleep regardless of age or dementia, the inclusion of wrist temperature as a complementary sensor to movement-based ones could enable their strengths to be leveraged and fused to support more accurate sleep detection.

More accurate sleep detection for older adults using wearable systems could support better short-term and long-term monitoring of sleep. Through daily sleep tracking, irregular sleep quality and patterns can be recognized by machine learning algorithms and then be reported to older adults or their health care providers. Such monitoring systems are especially useful for OAWD because the incidence of poor sleep is quite high in this population, and self-reporting of sleep is often difficult or impossible. Wearable technologies for sleep tracking and sleep problem identification could help to better understand older adults' sleep using objective measures, but this approach is only possible if the results are accurate. Accurate sleep data would enable the creation of tailored sleep management plans, enabling better sleep support for each individual, thus supporting better health outcomes and quality of life.

Limitations

This study had several limitations. First, only the Mi Band 2 was the only commercially available wearable device that was examined. As other commercially available smart wristbands collect other types of data (eg, PPG) and use different sleep detection algorithms, their performance will likely be different. These alternatives should be examined in healthy older adults and OAWD in future studies. Second, the sample size of this study was small. Future research with larger sample sizes is required to provide a deeper, comprehensive understanding that needs to support greater generalization of the results. Third, all our OAWD participants lived in the same LTC home. Data from other LTC homes as well as OAWD living in the community need to be investigated to determine how the environment plays a role in sleep detection and dementia. Finally, sleep onset and offset were estimated from sleep journals and/or actigraphy data. Using another method, such as computer vision, could estimate onset and offset more accurately as well as provide information about what was happening during sleep (eg, giving insight into repetitive movements).

Conclusions

This research adds evidence that wrist temperature can be used as an indicator of sleep status, including for OAWD and people with irregular sleep patterns. As illustrated by case studies from our data, suboptimal sleep detection performance by a commercial wristband was likely because of broken sleep patterns and body movements. As wearable technologies are increasingly being used by LTC homes and healthy older adults to track sleep and inform sleep management and support, this research suggests that caution should be used when interpreting sleep data when monitoring the sleep of older adults, particularly those living with dementia. This highlights the need for future research and development to create systems that better complement older adult populations. Future sleep monitoring wearable systems could consider adding a temperature sensor to capture the waist temperature as an extra indicator of sleep combined with motion-based data and machine learning algorithms of typical sleep patterns.

Acknowledgments

This study was funded by the principal investigator's Research Chair from the Schlegel-University of Waterloo Research Institute for Aging. JW was affiliated with the Department of Systems Design Engineering at the University of Waterloo at the time of the study and is currently affiliated with School of Computing and Information System at the University of Melbourne.

Conflicts of Interest

None declared.

References

1. Buxton OM, Marcelli E. Short and long sleep are positively associated with obesity, diabetes, hypertension, and cardiovascular disease among adults in the United States. Soc Sci Med 2010 Sep;71(5):1027-1036. [doi: 10.1016/j.socscimed.2010.05.041]
2. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003 Mar 15;26(2):117-126. [Medline: 12683469]
3. Walker MP, Stickgold R. Sleep, memory, and plasticity. Annu Rev Psychol 2006;57:139-166. [doi: 10.1146/annurev.psych.56.091103.070307] [Medline: 16318592]
4. Hudson AN, Van Dongen HP, Honn KA. Sleep deprivation, vigilant attention, and brain function: a review. Neuropsychopharmacol 2019 Jun 8;45(1):21-30. [doi: 10.1038/s41386-019-0432-6]
5. Ju YE, Lucey BP, Holtzman DM. Sleep and Alzheimer disease pathology--a bidirectional relationship. Nat Rev Neurol 2014 Feb;10(2):115-119 [FREE Full text] [doi: 10.1038/nrneurol.2013.269] [Medline: 24366271]
6. Musiek ES, Xiong DD, Holtzman DM. Sleep, circadian rhythms, and the pathogenesis of Alzheimer disease. Exp Mol Med 2015 Mar 13;47(3):e148. [doi: 10.1038/emm.2014.121]
7. Cordone S, Annarumma L, Rossini PM, De Gennaro L. Sleep and β-Amyloid deposition in Alzheimer disease: insights on mechanisms and possible innovative treatments. Front Pharmacol 2019;10:695 [FREE Full text] [doi: 10.3389/fphar.2019.00695] [Medline: 31281257]
8. Drager LF, McEvoy RD, Barbe F, Lorenzi-Filho G, Redline S, INCOSACT Initiative (International Collaboration of Sleep Apnea Cardiovascular Trialists). Sleep apnea and cardiovascular disease: lessons from recent trials and need for team science. Circulation 2017 Nov 07;136(19):1840-1850 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.117.029400] [Medline: 29109195]
9. Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI, et al. The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. Eur Heart J 2019 Apr 07;40(14):1149-1157 [FREE Full text] [doi: 10.1093/eurheartj/ehy624] [Medline: 30376054]
10. Kushida CA, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. Sleep Med 2001 Sep;2(5):389-396. [doi: 10.1016/s1389-9457(00)00098-8] [Medline: 14592388]
11. Kingshott RN, Douglas NJ. The effect of in-laboratory polysomnography on sleep and objective daytime sleepiness. Sleep 2000 Dec 15;23(8):1109-1113. [Medline: 11145325]
12. Sadeh A, Hauri PJ, Kripke DF, Lavie P. The role of actigraphy in the evaluation of sleep disorders. Sleep 1995 May;18(4):288-302. [doi: 10.1093/sleep/18.4.288] [Medline: 7618029]
13. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011 Aug;15(4):259-267. [doi: 10.1016/j.smrv.2010.10.001] [Medline: 21237680]
14. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003 May 01;26(3):342-392. [Medline: 12749557]
15. Smith MT, McCrae CS, Cheung J, Martin JL, Harrod CG, Heald JL, et al. Use of actigraphy for the evaluation of sleep disorders and circadian rhythm sleep-wake disorders: an American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med 2018 Jul 15;14(7):1231-1237 [FREE Full text] [doi: 10.5664/jcsm.7230] [Medline: 29991437]
16. Refinetti R, Menaker M. The circadian rhythm of body temperature. Physiol Behav 1992 Mar;51(3):613-637. [doi: 10.1016/0031-9384(92)90188-8] [Medline: 1523238]
17. Sarabia JA, Rol MA, Mendiola P, Madrid JA. Circadian rhythm of wrist temperature in normal-living subjects A candidate of new index of the circadian system. Physiol Behav 2008 Nov 28;95(4):570-580. [doi: 10.1016/j.physbeh.2008.08.005] [Medline: 18761026]
18. Ortiz-Tudela E, Martinez-Nicolas A, Campos M, Rol MA, Madrid JA. A new integrated variable based on thermometry, actimetry and body position (TAP) to evaluate circadian system status in humans. PLoS Comput Biol 2010 Nov 11;6(11):e1000996 [FREE Full text] [doi: 10.1371/journal.pcbi.1000996] [Medline: 21085644]
19. Wei J, Zhang J, Boger J. What wrist temperature tells us when we sleep late: a new perspective of sleep health. In: Proceedings of the IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation

(SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). 2018 Presented at: IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI); Oct 8-12, 2018; Guangzhou, China p. A. [doi: 10.1109/SmartWorld.2018.00145]
20. Wei J, Boger J. You are how you sleep: personalized sleep monitoring based on wrist temperature and accelerometer data. In: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare - Demos and Posters. 2019 Presented at: 13th EAI International Conference on Pervasive Computing Technologies for Healthcare - Demos and Posters; June 11, 2019; Trento, Italy. [doi: 10.4108/eai.20-5-2019.2282879]
21. Guillodo E, Lemey C, Simonnet M, Walter M, Baca-García E, Masetti V, et al. Clinical applications of mobile health wearable–based sleep monitoring: systematic review. JMIR Mhealth Uhealth 2020 Apr 1;8(4):e10733. [doi: 10.2196/10733]
22. Madrid-Navarro CJ, Cuesta FJ, Escamilla-Sevilla F, Campos M, Ruiz Abellán F, Rol MA, et al. Validation of a device for the ambulatory monitoring of sleep patterns: a pilot study on Parkinson's disease. Front Neurol 2019;10:356 [FREE Full text] [doi: 10.3389/fneur.2019.00356] [Medline: 31031690]
23. Landry GJ, Best JR, Liu-Ambrose T. Measuring sleep quality in older adults: a comparison using subjective and objective methods. Front Aging Neurosci 2015;7:166 [FREE Full text] [doi: 10.3389/fnagi.2015.00166] [Medline: 26441633]
24. Potvin O, Lorrain D, Forget H, Dubé M, Grenier S, Préville M, et al. Sleep quality and 1-year incident cognitive impairment in community-dwelling older adults. Sleep 2012 Apr 01;35(4):491-499 [FREE Full text] [doi: 10.5665/sleep.1732] [Medline: 22467987]
25. Nebes RD, Buysse DJ, Halligan EM, Houck PR, Monk TH. Self-reported sleep quality predicts poor cognitive performance in healthy older adults. J Gerontol B Psychol Sci Soc Sci 2009 Mar;64(2):180-187 [FREE Full text] [doi: 10.1093/geronb/gbn037] [Medline: 19204069]
26. Wang H, Zhao Y, Yu L, Liu J, Zwetsloot IM, Cabrera J, et al. A personalized health monitoring system for community-dwelling elderly people in Hong Kong: design, implementation, and evaluation study. J Med Internet Res 2020 Sep 30;22(9):e19223 [FREE Full text] [doi: 10.2196/19223] [Medline: 32996887]
27. Jirong Y, Changquan H, Hongmei W, Bi-Rong D. Association of sleep quality and dementia among long-lived Chinese older adults. Age (Dordr) 2013 Aug;35(4):1423-1432 [FREE Full text] [doi: 10.1007/s11357-012-9432-8] [Medline: 22669593]
28. Kim JH, Duffy JF. Circadian rhythm sleep-wake disorders in older adults. Sleep Med Clin 2018 Mar;13(1):39-50. [doi: 10.1016/j.jsmc.2017.09.004] [Medline: 29412982]
29. Gulia KK, Kumar VM. Sleep disorders in the elderly: a growing challenge. Psychogeriatrics 2018 May;18(3):155-165. [doi: 10.1111/psyg.12319] [Medline: 29878472]
30. Shi L, Chen S, Ma M, Bao Y, Han Y, Wang Y, et al. Sleep disturbances increase the risk of dementia: a systematic review and meta-analysis. Sleep Med Rev 2018 Aug;40:4-16. [doi: 10.1016/j.smrv.2017.06.010]
31. Rongve A, Boeve BF, Aarsland D. Frequency and correlates of caregiver-reported sleep disturbances in a sample of persons with early dementia. J Am Geriatr Soc 2010 Mar;58(3):480-486. [doi: 10.1111/j.1532-5415.2010.02733.x] [Medline: 20398116]
32. Naismith SL, Glozier N, Burke D, Carter PE, Scott E, Hickie IB. Early intervention for cognitive decline: is there a role for multiple medical or behavioural interventions? Early Interv Psychiatry 2009 Feb;3(1):19-27. [doi: 10.1111/j.1751-7893.2008.00102.x] [Medline: 21352171]
33. Yaffe K, Blackwell T, Barnes DE, Ancoli-Israel S, Stone KL, Study of Osteoporotic Fractures Group. Preclinical cognitive decline and subsequent sleep disturbance in older women. Neurology 2007 Jul 17;69(3):237-242. [doi: 10.1212/01.wnl.0000265814.69163.da] [Medline: 17636060]
34. Leng Y, Musiek ES, Hu K, Cappuccio FP, Yaffe K. Association between circadian rhythms and neurodegenerative diseases. Lancet Neurol 2019 Mar;18(3):307-318 [FREE Full text] [doi: 10.1016/S1474-4422(18)30461-7] [Medline: 30784558]
35. Yu RP, Ellison NB, McCammon RJ, Langa KM. Mapping the two levels of digital divide: internet access and social network site adoption among older adults in the USA. Inf Commun Soc 2015 Nov 19;19(10):1445-1464. [doi: 10.1080/1369118x.2015.1109695]
36. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018 Jan;2(1):igy002 [FREE Full text] [doi: 10.1093/geroni/igy002] [Medline: 30480129]
37. Fitbit. URL: https://www.fitbit.com [accessed 2021-04-30]
38. Mi Band. URL: https://www.mi.com/global/miband/ [accessed 2021-04-30]
39. Garmin. URL: https://www.garmin.com [accessed 2021-04-30]
40. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. Clin Pharmacol Ther 2018 Jul;104(1):42-52 [FREE Full text] [doi: 10.1002/cpt.966] [Medline: 29205294]
41. Degroote L, Hamerlinck G, Poels K, Maher C, Crombez G, De Bourdeaudhuij I, et al. Low-cost consumer-based trackers to measure physical activity and sleep duration among adults in free-living conditions: validation study. JMIR Mhealth Uhealth 2020 May 19;8(5):e16674. [doi: 10.2196/16674]

42. Liu J, Zhao Y, Lai B, Wang H, Tsui KL. Wearable device heart rate and activity data in an unsupervised approach to personalized sleep monitoring: algorithm validation. JMIR Mhealth Uhealth 2020 Aug 05;8(8):e18370 [FREE Full text] [doi: 10.2196/18370] [Medline: 32755887]
43. Gibson RH, Gander PH. Monitoring the sleep patterns of people with dementia and their family carers in the community. Australas J Ageing 2019 Mar;38(1):47-51. [doi: 10.1111/ajag.12605] [Medline: 30525266]
44. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4(2):97-110. [Medline: 1027738]
45. 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. [Medline: 2748771]
46. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [Medline: 1798888]
47. Monk TH, Reynolds CF, Kupfer DJ, Buysse DJ, Coble PA, Hayes AJ, et al. The Pittsburgh sleep diary. J Sleep Res 1994;3:111-120 [FREE Full text] [Medline: 11537903]
48. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc 1992 Sep;40(9):922-935. [Medline: 1512391]
49. Martinez-Nicolas A, Madrid JA, García FJ, Campos M, Moreno-Casbas MT, Almaida-Pagán PF, et al. Circadian monitoring as an aging predictor. Sci Rep 2018 Oct 09;8(1):15027 [FREE Full text] [doi: 10.1038/s41598-018-33195-3] [Medline: 30301951]
50. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [Medline: 1202204]
51. Huang Y, Liu R, Wang Q, Van Someren EJ, Xu H, Zhou J. Age-associated difference in circadian sleep-wake and rest-activity rhythms. Physiol Behav 2002 Aug;76(4-5):597-603. [doi: 10.1016/s0031-9384(02)00733-3] [Medline: 12126998]
52. Hibi S, Yamaguchi Y, Umeda-Kameyama Y, Yamamoto H, Iijima K, Momose T, et al. The high frequency of periodic limb movements in patients with Lewy body dementia. J Psychiatr Res 2012 Dec;46(12):1590-1594. [doi: 10.1016/j.jpsychires.2012.07.007] [Medline: 22835911]
53. Pao WC, Boeve BF, Ferman TJ, Lin S, Smith GE, Knopman DS, et al. Polysomnographic findings in dementia with Lewy bodies. Neurologist 2013 Jan;19(1):1-6. [doi: 10.1097/NRL.0b013e31827c6bdd] [Medline: 23269098]
54. McCurry SM, Ancoli-Israel S. Sleep dysfunction in Alzheimer's disease and other Dementias. Curr Treat Options Neurol 2003 May;5(3):261-272. [doi: 10.1007/s11940-003-0017-9] [Medline: 12670415]
55. Kang SG, Kang JM, Ko K, Park S, Mariani S, Weng J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. J Psychosom Res 2017 Jun;97:38-44. [doi: 10.1016/j.jpsychores.2017.03.009] [Medline: 28606497]

Abbreviations

CBT: core body temperature
ESS: Epworth Sleepiness Scale
IS: interday stability
IV: intraday variability
LTC: long-term care
MMSE: Mini-Mental State Examination
OAWD: older adults living with dementia
PSG: polysomnography
PSQI: Pittsburgh Sleep Quality Index
PSW: personal support worker

Edited by L Buis; submitted 12.12.20; peer-reviewed by H Kazemi Arpanahi, V Stara; comments to author 24.01.21; revised version received 03.02.21; accepted 12.04.21; published 01.06.21.

Please cite as:
Wei J, Boger J
Sleep Detection for Younger Adults, Healthy Older Adults, and Older Adults Living With Dementia Using Wrist Temperature and Actigraphy: Prototype Testing and Case Study Analysis
JMIR Mhealth Uhealth 2021;9(6):e26462
URL: https://mhealth.jmir.org/2021/6/e26462
doi:10.2196/26462
PMID:34061038

Original Paper

Smartphone-Based VO2max Measurement With Heart Snapshot in Clinical and Real-world Settings With a Diverse Population: Validation Study

Dan E Webster[1*], PhD; Meghasyam Tummalacherla[1*], MS; Michael Higgins[2], MS, XT; David Wing[2], CPT, MS, CBDT; Euan Ashley[3], MD, PhD; Valerie E Kelly[4], PT, PhD; Michael V McConnell[5,6], MD, MSEE; Evan D Muse[7], MD; Jeffrey E Olgin[8], MD; Lara M Mangravite[1], PhD; Job Godino[2,7], PhD; Michael R Kellen[1*], PhD; Larsson Omberg[1*], PhD

[1]Sage Bionetworks, Seattle, WA, United States

[2]Exercise and Physical Activity Resource Center, University of California at San Diego, San Diego, CA, United States

[3]Department of Genetics, Stanford University School of Medicine, Stanford, CA, United States

[4]Department of Rehabilitation Medicine, University of Washington, Seattle, WA, United States

[5]Stanford University School of Medicine, Stanford, CA, United States

[6]Google Health, Palo Alto, CA, United States

[7]Scripps Research Translational Institute and Scripps Clinic, La Jolla, CA, United States

[8]Division of Cardiology and the Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Larsson Omberg, PhD
Sage Bionetworks
2901 3rd Ave #330
Seattle, WA
United States
Phone: 1 206 928 8250
Email: larsson.omberg@sagebionetworks.org

Abstract

Background: Maximal oxygen consumption (VO_2max) is one of the most predictive biometrics for cardiovascular health and overall mortality. However, VO_2max is rarely measured in large-scale research studies or routine clinical care because of the high cost, participant burden, and requirement for specialized equipment and staff.

Objective: To overcome the limitations of clinical VO_2max measurement, we aim to develop a digital VO_2max estimation protocol that can be self-administered remotely using only the sensors within a smartphone. We also aim to validate this measure within a broadly representative population across a spectrum of smartphone devices.

Methods: Two smartphone-based VO_2max estimation protocols were developed: a 12-minute run test (12-MRT) based on distance measured by GPS and a 3-minute step test (3-MST) based on heart rate recovery measured by a camera. In a 101-person cohort, balanced across age deciles and sex, participants completed a gold standard treadmill-based VO_2max measurement, two silver standard clinical protocols, and the smartphone-based 12-MRT and 3-MST protocols in the clinic and at home. In a separate 120-participant cohort, the video-based heart rate measurement underlying the 3-MST was measured for accuracy in individuals across the spectrum skin tones while using 8 different smartphones ranging in cost from US $99 to US $999.

Results: When compared with gold standard VO_2max testing, Lin concordance was p_c=0.66 for 12-MRT and p_c=0.61 for 3-MST. However, in remote settings, the 12-MRT was significantly less concordant with the gold standard (p_c=0.25) compared with the 3-MST (p_c=0.61), although both had high test-retest reliability (12-MRT intraclass correlation coefficient=0.88; 3-MST intraclass correlation coefficient=0.86). On the basis of the finding that 3-MST concordance was generalizable to remote settings whereas 12-MRT was not, the video-based heart rate measure within the 3-MST was selected for further investigation. Heart rate

measurements in any of the combinations of the six Fitzpatrick skin tones and 8 smartphones resulted in a concordance of p_c≥0.81. Performance did not correlate with device cost, with all phones selling under US $200 performing better than p_c>0.92.

Conclusions: These findings demonstrate the importance of validating mobile health measures in the real world across a diverse cohort and spectrum of hardware. The 3-MST protocol, termed as *heart snapshot*, measured VO_2max with similar accuracy to supervised in-clinic tests such as the Tecumseh (p_c=0.94) protocol, while also generalizing to remote and unsupervised measurements. *Heart snapshot* measurements demonstrated fidelity across demographic variation in age and sex, across diverse skin pigmentation, and between various iOS and Android phone configurations. This software is freely available for all validation data and analysis code.

(JMIR Mhealth Uhealth 2021;9(6):e26006) doi:10.2196/26006

KEYWORDS

VO2max; heart rate; digital health; real-world data; cardiorespiratory fitness; remote monitoring; mobile phone; smartphone; validation

Introduction

Background

Expanding access to precision medicine will increasingly require patient biometrics to be measured in remote care settings. Traditionally, cardiovascular health has been assessed using risk scores such as the Framingham Risk Score [1], Reynolds Risk Score [2], Qrisk [3], and others, which integrate multiple factors including demographic data, comorbidities, and biometrics paired with imaging-based assessments measuring vascular blockage and blood flow in higher-risk and symptomatic individuals. Although these factors have a clear correlation with cardiovascular health, their inclusion in integrative risk calculations was promoted in part because they can be rapidly evaluated across many individuals. However, one of the most predictive biometrics for cardiovascular health [4] and overall mortality [5], maximal oxygen consumption (VO_2max), is typically not incorporated in these risk calculators because of the high cost, participant burden, and specialized equipment and staff needed to obtain this measurement [6,7].

Cardiorespiratory fitness, as measured by VO_2max, represents the integrated function of physiological systems involved in transporting oxygen from the atmosphere to the skeletal muscles to perform physical work. Existing gold standard techniques for measuring VO_2max are based on protocols that use exercise on a treadmill or stationary bicycle paired with the direct measurement of oxygen consumption at various workloads, including maximal exertion [8,9]. However, the requirement to exercise at the maximal aerobic threshold limits deployment in some populations for safety reasons, and the need for specialized equipment and personnel has prohibited widespread adoption of VO_2max testing in research and clinical settings.

Objectives

Limitations of gold standard VO_2max measurements have led to the development of numerous ”silver standard” [10] VO_2max estimation protocols that rely on simpler equipment or submaximal levels of exertion. These protocols trade off measurement accuracy for ease of deployment in a wider range of settings and for populations with differing levels of capacity [11]. However, these protocols were typically developed and validated in small, homogeneous populations, and some subsequent validation studies have been criticized for demonstrating participant selection bias [12]. To overcome these limitations, we aim to develop a digital VO_2max estimation protocol that could be self-administered remotely using only the sensors within a smartphone, and we also aim to validate this measure within a broadly representative population. Previous efforts have used a smartphone-based approach to measure VO_2max, but these validation studies are rarely conducted outside of clinical settings [13]. Therefore, we aim to validate our measurements in remote, unsupervised real-world settings.

Methods

Development of Smartphone Sensor–Based Measurements of VO_2max

Two silver standard VO_2max estimation protocols were chosen as the basis for developing the smartphone tests. The first is the Cooper protocol [14], comprising a 12-minute run test (12-MRT), where individuals cover as much distance as possible in 12 minutes on a flat course. The Cooper protocol estimates VO_2max from the total distance traveled during the 12 minutes. The other is the Tecumseh protocol [15], which comprises a 3-minute step test (3-MST), where individuals step up and down an 8-inch step at a constant rate for 3 minutes. In the 3-MST protocol, VO_2max was estimated from the heart rate measurements during the recovery period. In adopting these protocols for smartphones, we developed self-guided instructions with GPS to record distance during the 12-MRT and a smartphone camera to record heart rate during recovery for the 3-MST (Multimedia Appendix 1).

VO_2max Validation Cohort Procedures and Measures

All study procedures were approved by the University of California, San Diego (UCSD) Institutional Review Board (approval number 171815). All participants provided written informed consent and attended two in-person study visits at the Exercise and Physical Activity Resource Center (EPARC).

A convenience sample of 101 adults aged between 20 and 79 years was recruited, largely balanced across age deciles and sex (Multimedia Appendix 2). Potential participants were contacted by trained EPARC staff via email or telephone and they

underwent a screening to ascertain their eligibility. Participants were included if they were (1) able to consent and participate in the study using English; (2) aged between 20 and 79 years; (3) willing and able to attend two in-person study visits that included either a VO_2max test or a 12-MRT and a 3-MST within a 2-week period; (4) willing and able to undertake up to three 12-MRT and 3-MST at home over a 2-week period; (5) willing and able to download the smartphone app developed to measure cardiorespiratory fitness on a compatible Android or iOS device and use it during all tests over a 2-week period; and (6) willing and able to download the Fitbit smartphone app on a compatible Android or iOS device and connect and wear a study-provided Fitbit Charge 2 during all tests over a 2-week period. Participants were excluded if they (1) were >12 weeks pregnant; (2) had a heart or cardiovascular condition, including coronary artery disease, congestive heart failure, diagnosed abnormality of heart rhythm, atrial fibrillation, and/or a history of myocardial infarction; (3) required the use of an external device to assist heart rhythm (eg, a pacemaker); (4) had a serious respiratory disease, including chronic obstructive pulmonary disease, exercise-induced asthma, and/or pulmonary high blood pressure; (5) required use of supplemental oxygen; (6) required use of a beta-blocker or other medications known to alter heart rate; and (7) answered "yes" to one or more questions in the American College of Sports Medicine's Physical Activity Readiness Questionnaire and/or reported two or more risk factors for exercise testing and did not receive subsequent medical clearance. The Physical Activity Readiness Questionnaire is a widely accepted tool used to assess an individual's fitness for tests involving cardiovascular exercise [16].

Upon completion of the telephone screening (and, if necessary, receipt of medical clearance), potential participants were scheduled to attend the first testing session at the UCSD. They were asked to report to the testing session well hydrated and in an athletic attire. Participants were guided through the process of downloading and installing the smartphone app developed to measure cardiorespiratory fitness, as well as the Fitbit smartphone app, and they were fitted with a wrist-worn Fitbit Charge 2 according to the manufacturer's recommendations. Participants were asked to provide their age, sex at birth, ethnicity, and race. Weight (to the nearest 0.1 kg) and height (to the nearest 0.1 cm) were measured using a calibrated digital scale and stadiometer (Seca 703, Seca GmbH & Co. KG). Both weight and height were measured with participants wearing lightweight clothes without shoes, and two separate measurements were averaged (if weight or height measurements differed by more than 1%, then a third measure was taken, and the average of the two measures that differed by less than 0.2 kg or 0.5 cm, respectively, was used).

At the first testing session, participants either undertook a VO_2max test or an in-clinic 3-MST and 12-MRT. A randomization procedure implemented before the scheduling of the first testing session determined which test procedure participants undertook during the first testing session. The participants were then expected to complete the other test procedures during the second testing session.

Treadmill-Based Gold Standard VO_2max Measurement

Participants completed a maximal graded exercise test on a Woodway 4Front treadmill (Woodway) calibrated monthly for accuracy of speed and grade. The maximal graded exercise test protocol began with a warm-up at a self-selected pace on a treadmill for 5 to 10 minutes. During the warm-up, EPARC staff explained how to use the Borg Rating of Perceived Exertion scale and reminded participants that they were expected to achieve their maximal level of exertion [17].

The participants were then equipped with a breath mask that covered the nose and mouth (KORR Medical Technologies) and a Bluetooth-enabled heart rate monitor worn on the chest (Garmin). The preprogrammed treadmill protocol began with the participants running at 5 mph with a 0% incline for 3 minutes. The workload was then increased by approximately 0.75, which is the metabolic equivalent of tasks every minute. This was achieved via an increase in speed (0.5 mph per min) each minute until the participant was 0.5 mph above their self-determined comfortable speed or until a maximal speed of 9 mph was reached. If the participant's capacity allowed them to continue beyond this upper speed limit before reaching volitional fatigue, then the treadmill speed was kept constant, but the grade (ie, incline) of the treadmill was increased by 1% each minute until volitional fatigue was reached. The Borg Rating of Perceived Exertion scale was assessed during the final 10 seconds of each minute, and the protocol continued until the participant signaled to stop (ie, indication of volitional fatigue). Upon indication of volitional fatigue, the treadmill was immediately slowed to 2 mph, and participants were encouraged to walk until completely recovered. Breath-by-breath oxygen uptake (VO_2) was continuously measured using an indirect calorimeter (COSMED) that was calibrated for gas volume and fractional composition immediately (ie, <30 min) before the start of the maximal graded exercise test protocol.

Tecumseh Test (3-MST) and Cooper Test (12-MRT) In-Clinic Procedure

All participants were fitted with a chest-worn heart rate monitor (Polar) that was used for real-time monitoring by trained EPARC staff throughout both the 12-MRT and 3-MST. For the 3-MST, participants were instructed to step up and down from a single step 8 inches in height at a rate of 24 steps per minute for 3 minutes [18]. The cadence of stepping was monitored by trained EPARC staff. The radial pulse was measured from the 31st second to the 60th second after 3 minutes of stepping. Upon completion of the test, the participants were asked to sit in a chair and rest. After a minimum of 10 minutes of rest, participants completed a 5-minute self-determined light intensity warm-up. They were then instructed to cover as much distance as possible on a flat 400-m track for 12 minutes. The distance traveled was measured after 12 minutes [14].

Distance Estimation Using Privacy-Preserving GPS Data

The distance recorded by the smartphone during the 12-MRT was validated against the actual distance. The smartphone recorded displacement information sampled at 1 Hz, which

consists of relative location measurements, that is, the change in location with regard to the last recorded measurement. The iPhones (Apple Inc) measured displacement in meters whereas the Android smartphones measured relative changes in latitude and longitude, requiring an estimate of the absolute latitude and longitude to be added back into the measurements to obtain an accurate estimate of distance.

The first distance estimation method entailed summing the Euclidean distances between subsequent GPS points. As GPS measurements have a range error dependent on atmospheric effects and numerical errors, a second method was used to compute the distance after smoothing the trajectory of the GPS path using a Savitzky-Golay smoothing filter.

Camera-Based Heart Rate Estimation

Blood flow through the fingertip was measured through video with a rear-facing camera while the flash was on. The resting heart rate was captured with 20 seconds of recording, whereas the 3-MST required 60 seconds of recording. During the capture, we found it was important to fix the focal length to infinity, turn off any high dynamic range settings (if applicable), and set the frame rate to 60 Hz if possible, and if not, the default highest allowed by the phone. We did not record the video in order to preserve privacy associated with the inadvertent capture of identifiable objects in the frame before covering the lens with the finger, but instead summarized each video frame to the mean of all pixel intensities per color channel in the red, green, blue space.

These intensities yielded three time series, one for each color. These time series were filtered and mean-centered before being split into shorter 10-second windows. By assuming a periodic signal for these windows, the autocorrelation function (ACF) was used to estimate the period by finding the peaks and their corresponding lags. The relative magnitude of the peaks to the maxima of the ACF was used to generate a confidence score, which quantifies the extent to which the signal is periodic or if the peak at the fundamental frequency (ie, the peak with the highest magnitude) is a spurious peak. The ACF is calculated over a 10-second window, as this provides sufficient heart beat observations postprocessing to estimate heart rates ranging from 45 to 210 bpm.

To filter potentially spurious peaks, a magnitude threshold relative to the magnitude of the peak at zero lag was used. The confidence score was calculated as the ratio of the magnitude of the peak corresponding to the fundamental frequency to the next peak. The confidence score is an indicator of the periodicity of the signal, a property indicative of the heart rate signal in a short finite time window. The different color channels were merged by choosing the heart rate estimate from the channel (red or green) that had the maximum confidence score within a given window.

Estimation of VO_2max

3-Minute Step Test

Multiple formulas for predicting VO_2max from the Tecumseh step test and its variations have been developed [15]; here, we used the following established by Milligan [19]:

where HB3060 is the number of beats between *30* and *60* seconds after the step test, *age* is the age of the subject, and *sex* is *0* if male and *1* if female.

12-Minute Run Test

VO_2max for the 12-MRT is estimated from the following formula, where d_{12} is the distance covered in meters [14]:

Heart Rate Calibration Study Procedures and Measures

All study procedures were approved by the UCSD Institutional Review Board (approval number 181820). All participants provided written informed consent and attended one in-person study visit at the EPARC.

A convenience sample of 120 adults, aged 18-65 years, of six different skin types were asked to participate in this study. We aimed to recruit an equal ratio of male and female participants, as well as an equal number of participants with each skin type, as determined by the Fitzpatrick scale. Participants were included if they were (1) able to consent and participate in the study in English and (2) aged between 18 and 65 years. Participants were excluded if they had (1) peripheral neuropathy or (2) tattoos or scarring at the measurement site (index finger and/or wrist). Potential participants were contacted by trained EPARC staff via email or telephone, and they were asked to complete the screening to ascertain their eligibility.

To establish the Fitzpatrick skin type of the cohort during recruitment, participants were asked to self-assess their Fitzpatrick skin type based on visual comparison with images of well-known celebrities with diverse pigmentation levels. As self-assessment of skin type can have variable accuracy [20,21], spectrocolorimetry was also used as an objective standard [22]. Spectrocolorimetry measurements were performed on the underside of the jaw using Pantone RM200QC. To calculate pigmentation in the individual typology angle color space, the L^* and b^* parameters from the spectrocolorimetry measurements were used according to the formula:

$$\text{individual typology angle} = [\arctan((L^*-50)/b^*)] \times 180/3.14159 \quad \textbf{(1)}$$

Using this formula, skin color types can be classified into six groups, ranging from very light to dark skin: very light>55°>light>41°>intermediate>28°>tan>10°>brown>−30°>dark [22].

Upon completion of the telephone screening, potential participants were scheduled to attend the first testing session at the UCSD. Participants were asked to provide their age, sex at birth, ethnicity, and race. All participants were fitted with a chest-worn heart rate monitor that was used for real-time monitoring by trained EPARC staff throughout testing. Heart rate was also monitored using a finger-based pulse oximeter (Nonin Medical, Inc). The finger-based pulse oximeter was attached to the participants' index finger, and the time was

synchronized between the computer and the device. Trained research staff visually confirmed that the photoplethysmograph was reading accurately before starting measurements on smartphone devices.

Participants were then given the first of 8 smartphones: Huawei Mate SE, LG Stylo 4, Moto G6 Play, Samsung Galaxy J7, Samsung Galaxy S9+, iPhone8+, iPhoneSE, and iPhoneXS. They were instructed by trained research staff to stand still and gently cover the camera and flash on the back of the smartphone with their fingertip, as their heart rate was captured by our preloaded smartphone app. The time on the Polar app was recorded at the time the measurement began on the smartphone app. Measurements with each smartphone lasted 60 seconds in duration. Processed data from the finger-based pulse oximeter were parsed and transformed with custom scripts to generate continuous photoplethysmography data in a format suitable for comparison with the heart rates from the phones.

Statistical Analysis

Demographic data were described using univariate summary statistics (eg, proportions, means, and SDs). Test validity for heart rate estimates and VO_2max was visualized using Bland-Altman plots [23] and compared using the Lin concordance index [24]. The heart rate errors were also compared using percent error. Analyses were performed in both R and Python.

iOS and Android Heart Snapshot Software Modules

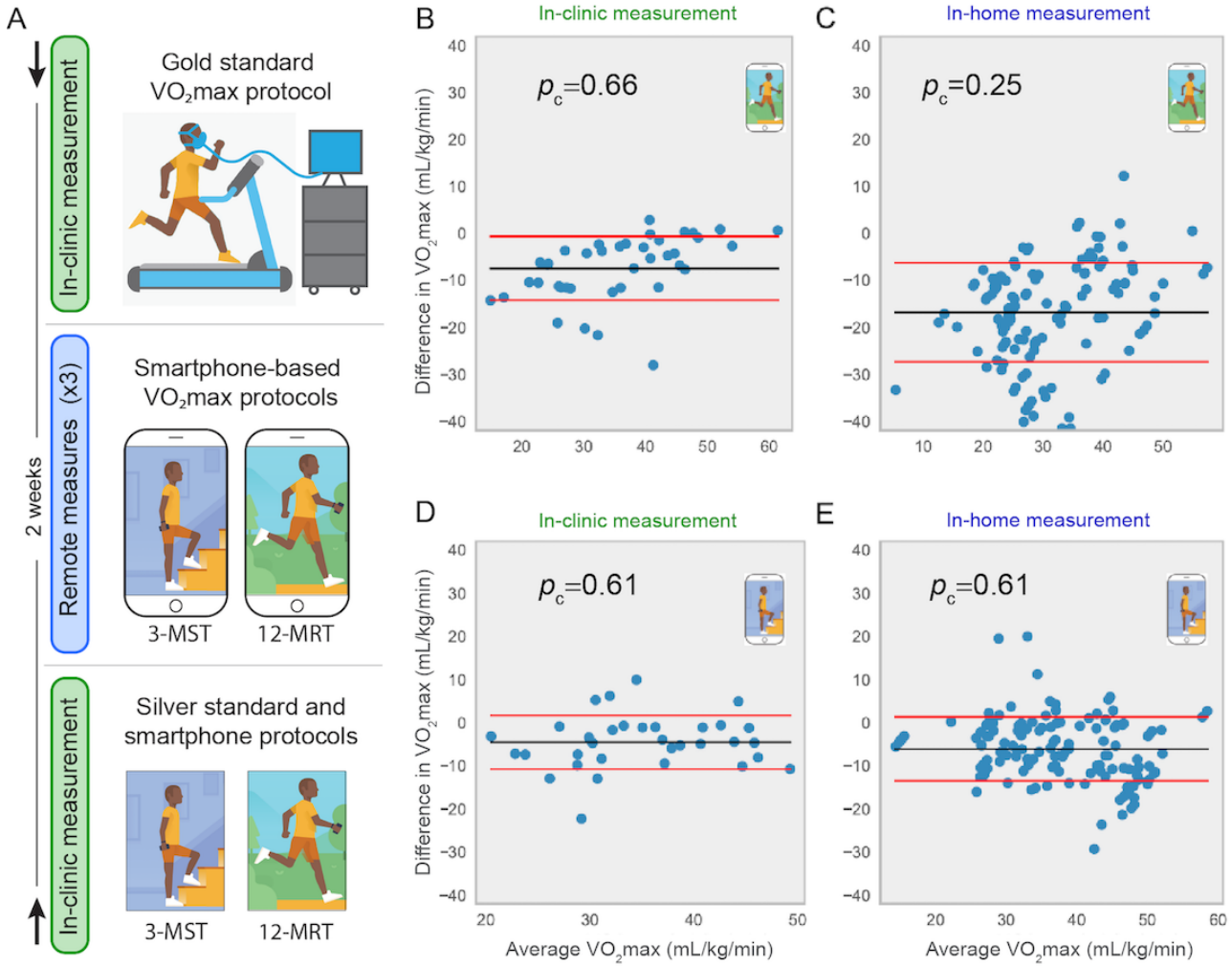
The code for the heart snapshot modules and sample Android [25] and iOS [26] apps are available under an open-source license.

Results

Validation in a Clinical Setting

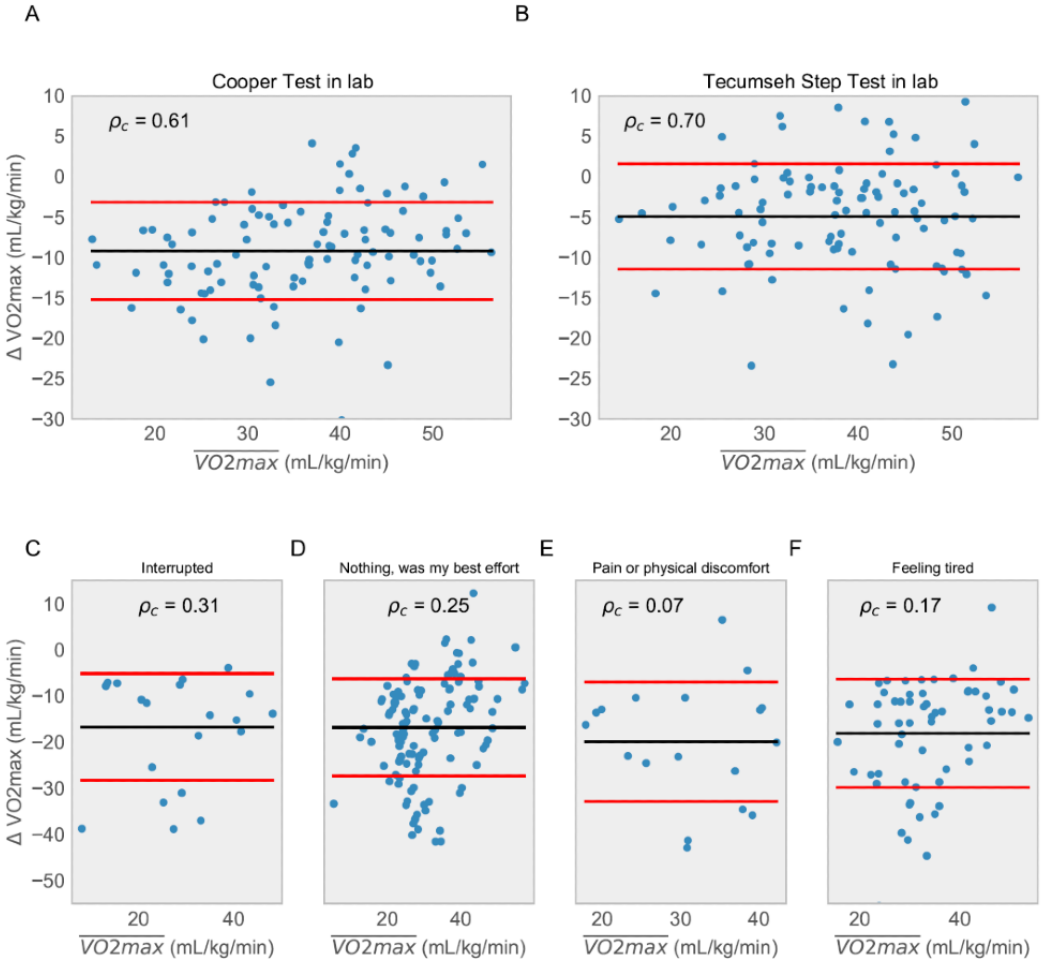
To assess the validity of the 3-MST and 12-MRT smartphone measurements, gold standard VO_2max treadmill testing was performed with 101 participants distributed across age deciles 20-80 years. Every participant also performed the silver standard and smartphone 12-MRT and 3-MST protocols in the clinic, with three instances of each smartphone protocol performed over 2 weeks without supervision in the participant's home environment (Figure 1).

Figure 1. Validation protocol and primary results of validation. (A) Participants in the study were randomized into two groups. The first group (denoted by the downward-facing arrow at top) performed a gold standard VO_2max protocol and received training on day 1. The second group performed the two silver standard protocols concurrently with the smartphone protocols on day 1 (denoted by the upward-facing arrow at bottom). Both groups then performed the two smartphone protocols remotely up to three times during a 2-week period. (B) to (E) show Bland-Altman plots comparing the gold standard VO_2max with smartphone measures from: (B) 12-MRT performed in clinic, (C) 12-MRT performed remotely (up to 3 repeats per participant), (D) 3-MST in clinic, and (E) 3-MST remotely. VO_2max: maximal oxygen consumption; 3-MST: 3-minute step test; 12-MRT: 12-minute run test.



The in-clinic 12-MRT distance was measured on a 400-m track and by the smartphone GPS. The in-clinic heart rate was measured via radial pulse by trained research staff, a chest-worn Polar heart monitor, a wrist-worn Fitbit Charge 2, and a smartphone camera with the flash activated. Comparisons between the gold standard, silver standard, and smartphone-based protocols for VO_2max estimation were performed using Bland-Altman analysis [23] and the Lin concordance index (p_c). The concordance between gold standard VO_2max and the silver standard Cooper protocol (p_c=0.61; Figure 2) and the silver standard Tecumseh protocol (p_c=0.70; Figure 2) were in line with previously published results [27-29]. Concordance of smartphone-based protocols with gold standard VO_2max testing was p_c=0.66 for the 12-MRT (Figure 1) and p_c=0.61 for the 3-MST (Figure 1). The concordance of smartphone-based protocols with silver standard protocols was p_c=0.96 for the 12-MRT and p_c=0.94 for the 3-MST. These results demonstrate that the smartphone-based protocols fall short of recapitulating gold standard VO_2max testing but are highly concordant with validated silver standard VO_2max estimation protocols in a laboratory setting.

Figure 2. Comparison of in-clinic performance of silver standard protocols relative to the gold standard for (A) 12-minute run test (12-MRT) and (B) 3-minute step test. For each plot, we are showing the difference between the ground truth maximal oxygen consumption measurement and measurements obtained using the distance run around a track (for A) and heart rate via radial pulse measured by trained research staff (for B) as per Tecumseh protocol. This distance was also measured using GPS and heart rate was measured using a chest strap and Fitbit. The concordance between distance measured around the track and measured using the GPS in the phone was 0.96. (C) to (F) show the concordance of the 12-MRT test for different values of self-reported effort. VO_2max: maximal oxygen consumption.



Validation in a Remote Setting

To investigate whether the concordance of in-clinic measurements would generalize to remote and unsupervised settings, the smartphone protocols were also performed up to three times at home by each participant. We observed an approximately equal test-retest reliability between the two tests (3-MST intraclass correlation coefficient=0.86; 12-MRT intraclass correlation coefficient=0.88). However, although the 3-MST translated well to an unsupervised setting (ρ_c=0.61; Figure 1), the 12-MRT demonstrated a pronounced drop in concordance (ρ_c=0.25; Figure 1), despite a highly accurate distance measurement from the smartphone (ρ_c=0.96) based on comparisons made in a clinical setting.

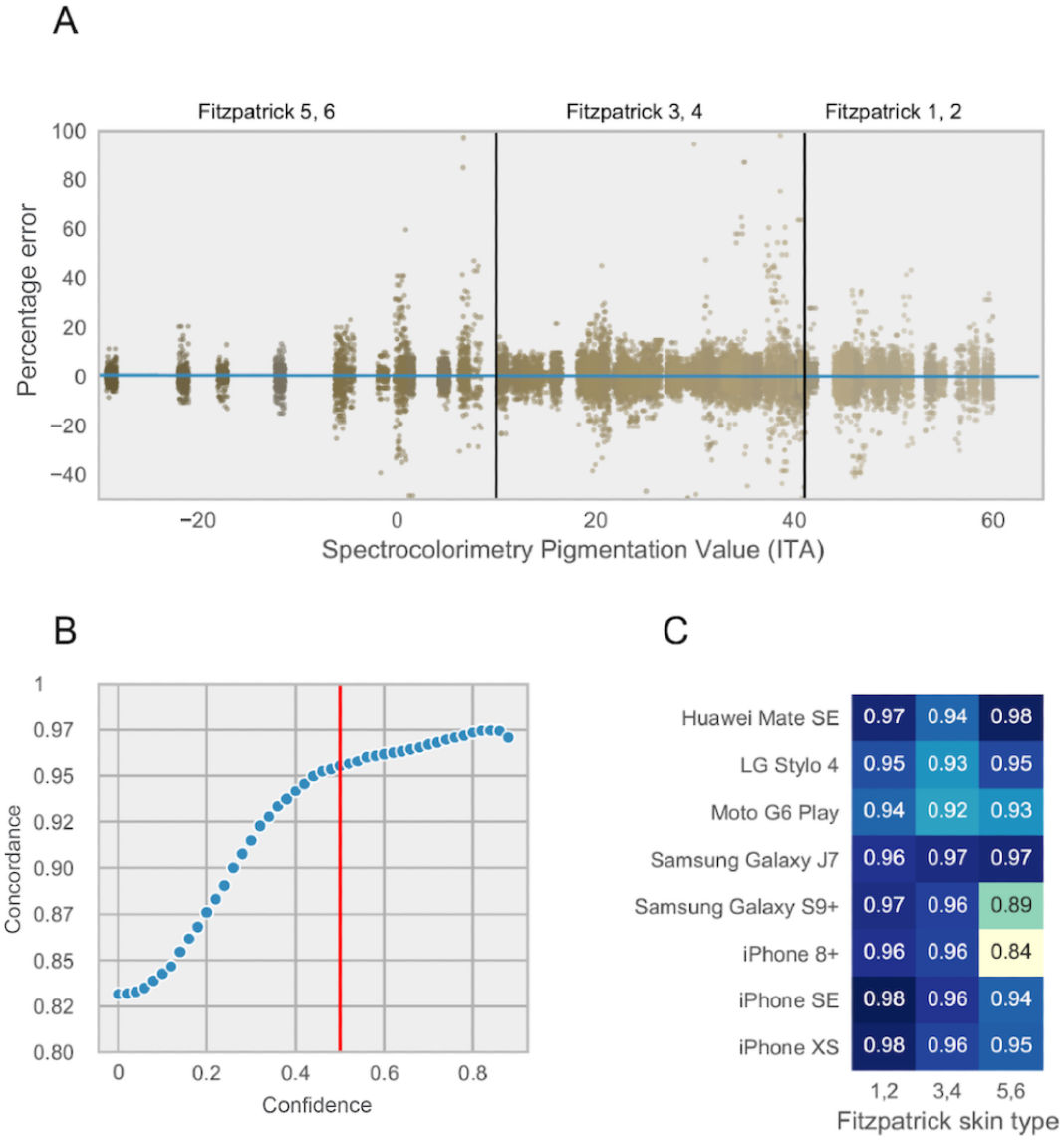
As the 12-MRT is dependent on maximal effort, participants were surveyed directly after their run about their performance. In 63.4% (137/216) of runs performed remotely, participants reported the run to be "their best effort." Therefore, only 137 runs were used to estimate VO_2max in our analysis. Figure 2 captures the results of all 216 runs subdivided by self-reported effort. Although the context-dependent failure of the 12-MRT in remote settings may be attributable to many factors, this result highlights the importance of both clinical and unsupervised real-world evidence for the validation of novel digital health measurement modalities.

Calibration of the Heart Snapshot Measurement for a Diverse Audience

The smartphone-based 3-MST protocol, hereafter referred to as *heart snapshot*, was generalizable between clinical and remote assessments and was robust over a large range of fitness levels. Maintaining high concordance during unsupervised measurements is necessary to achieve the scale intended for the targeted 1 million participants in the *All of Us* Research Program (AoURP) [30], which will use a "bring-your-own-device" strategy for remote self-measurement of VO_2max. AoURP also aims to recruit a study population matching the full demographic diversity of the United States, emphasizing the inclusion of groups often underrepresented in biomedical research, such as ethnic and racial minorities. As prior studies have shown differing results as to whether optical techniques for heart rate detection (photoplethysmography) can be demographically biased [31,32], we aimed to investigate any differences in *heart snapshot* accuracy across variations in skin tone. A follow-up calibration study for heart rate measurements was conducted

with 120 participants distributed approximately evenly across defined Fitzpatrick skin types [33], using 8 different smartphones (3 iPhones and 5 Android smartphones ranging in cost from US $99 to US $999 at the time of writing). These phones were chosen to be representative of different operating systems, quality of sensors, processing speed, and camera configuration. Importantly, we observed no significant difference in heart rate measurement accuracy between categorical Fitzpatrick skin types or systematic measurement error proportional to skin color at either end of the spectrum (Figure 3).

Figure 3. Validation of heart rate measurements across different skin tones and hardware configurations in the calibration study. (A) Percent error in heart rate estimation from ground truth as a function of different colors captured by spectrocolorimetry under the jaw. Each dot represents a 10 second window of heart rate in one individual. (B) Distribution of concordance between heart rate using pulse oximetry and smartphone as the confidence cutoff is changed. Red line represents the chosen cutoff used for analysis. (C) Concordance as a function of smartphone models and Fitzpatrick skin tones. ITA: individual typology angle.



Internal Quality Control Procedures for Heart Snapshot

To facilitate quality control of the measurements across different smartphones, a confidence score was developed to provide a readout of the quality of the heart rate measurements. This confidence score is derived from the ACF of the heart rate signal across 10-second measurement windows. Using the calibration study results, a balance between the quality of measurements was weighed against the loss of data by choosing a filtration cutoff at a confidence level of ≥0.5. This resulted in a p_c=0.95, in the calibration cohort between a pulse oximetry pulse measurement and the camera-estimated heart rate (Figure 3). In selecting this confidence score as a cutoff, we observed that 80.41% (28,032/34,859) of all measurement windows were retained in this calibration cohort (Figure 4). The same cutoff was used in the validation of *heart snapshot* against gold standard VO_2max, where the heart rate concordance with a chest-worn Polar heart monitor was p_c=0.95 and p_c=0.83 when compared with a wrist-worn Fitbit Charge 2, both at home and in the clinic. This can be compared with p_c=0.92 between Polar and Fitbit Charge 2 (Figure 5).

Figure 4. Effect of different confidence cutoff on the amount of missing data from the calibration study. (A) Distribution of best confidence across red and green channels in the calibrations study and (B) percent of the 10 second windows that are filtered out at different cutoffs of the confidence score. The cutoff used in the analysis is 0.5 marked by the red line.

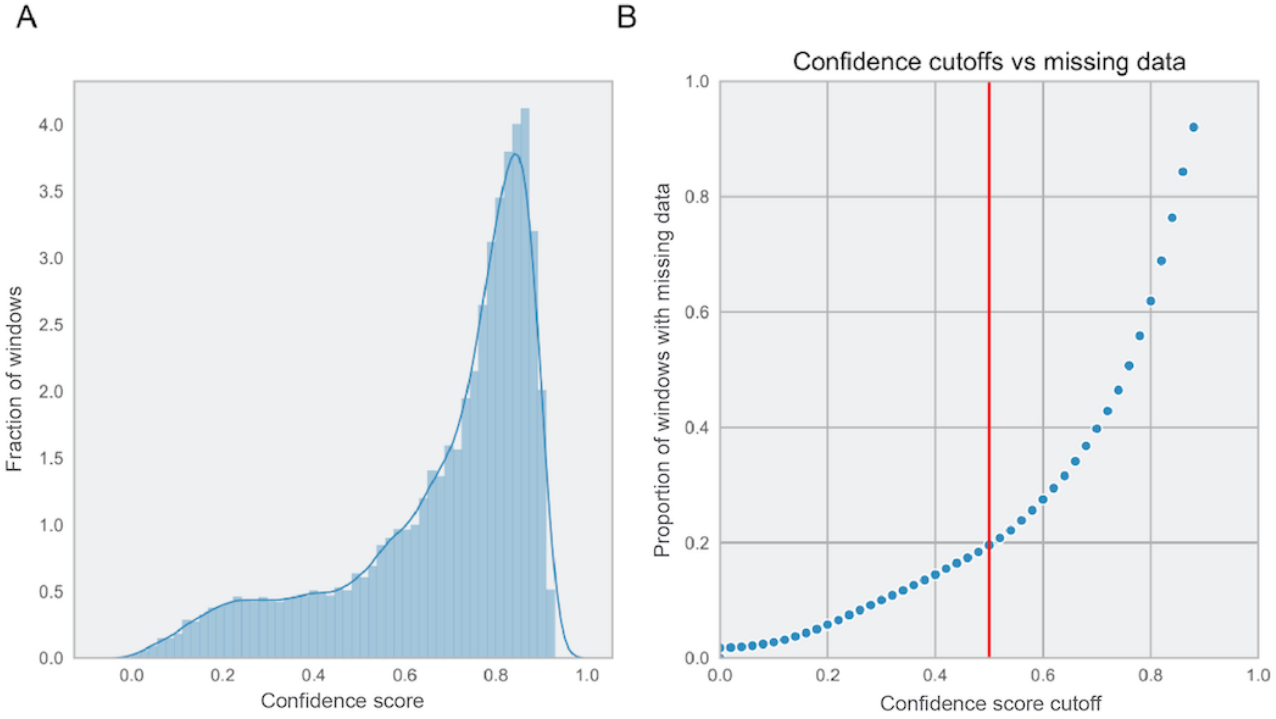
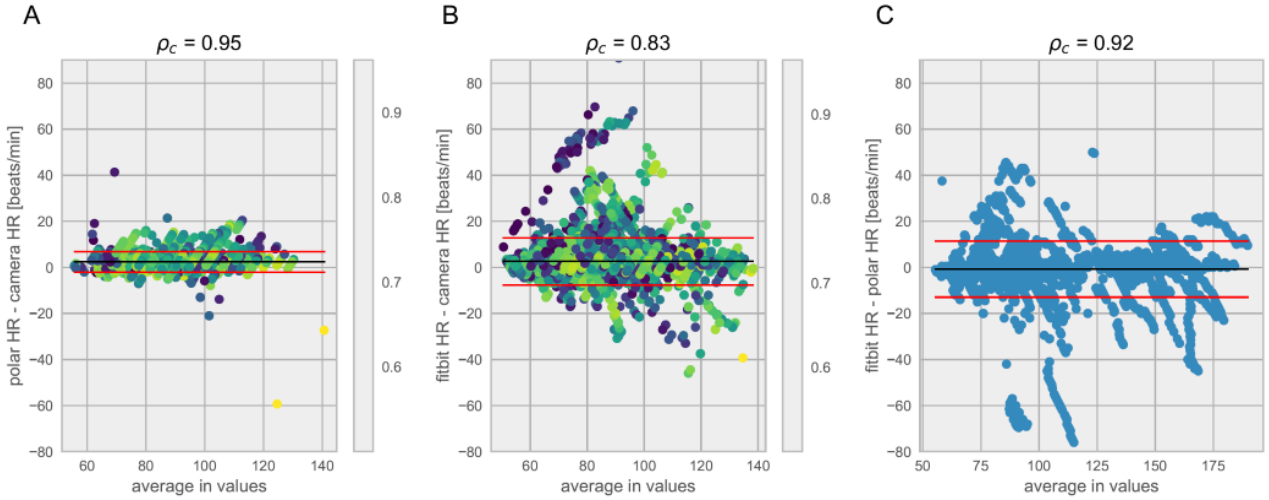


Figure 5. Bland-Altman analysis comparing heart rate measurements in the validation study using data collected during the Tecumseh tests. In the validation cohort, participants used multiple ways of collecting heart rate. The method being tested, the smartphone camera, was compared to: (A) a Polar chest strap (considered a gold standard) while in the clinic when both were used and (B) a Fitbit worn during the entirety of the study. (C) We also compared the Polar strap to the Fitbit for all time that both were worn. HR: heart rate.



Taken together, *heart snapshot* heart rate measurements in any of the combinations of the Fitzpatrick skin tones and 8 smartphones used in the calibration study resulted in a concordance greater than or equal to p_c=0.84 (Figure 3), which is in line with previous smartphone-based modalities for heart rate monitoring [34]. Importantly, performance did not correlate with device cost, with all phones selling for under US $200 performing better than p_c>0.92 for any skin tone.

Discussion

Principal Findings

In summary, *heart snapshot* measured VO_2max with similar accuracy to supervised, in-clinic tests such as the Tecumseh or Cooper protocols, while also generalizing to remote and unsupervised measurements. *Heart snapshot* measurements demonstrated fidelity across demographic variation in age and sex, across diverse skin pigmentation, and between various iOS and Android phone configurations.

The results from our validation study performed in unsupervised, remote environments showed that heart snapshot, which is based

on a 3-MST protocol, generalized to real-world settings but the 12-MRT protocol did not. Although it is difficult to definitively determine the reason for the poor concordance of 12-MRT, we suspect that this might be attributed to the Hawthorne effect, where people perform better when they are under constant observation at a track. It could also be purely environmental, where traffic, hills, and distractions impede uninterrupted running. This indicates the importance of testing and validating digital health measures in a representative setting.

Limitations

An important limitation of this study is that we did not include any individuals in our cohort with a known irregular heart rhythm, so we cannot extend our claims of validity to that population. Similarly, although we made efforts to include individuals across different age deciles, we focused solely on adults (aged over 18 years) and our age decile from 60 to 70 years did not include participants older than 65 years. This work was limited to the biometrics of resting heart rate and VO_2max, but using the same technology could also be extended to measure heart rate recovery in minute intervals after exertion, which would provide a valuable biometric that has been associated with prediction of overall mortality [30]. *Heart snapshot* attempts to maximize concordance with gold standard methods for estimating VO_2max, but it is worth noting that this analysis used an existing validated algorithm [27] that was based on in-clinic procedures and measurement tools. *Heart snapshot* could become more personalized than traditional protocols, for example, adapting to a participant's maximum step cadence as measured by smartphone accelerometry. Further concordance with gold standard measures may be achieved by optimizing the parameters of the traditional algorithm or including new variables, but this will require a distinct cohort to test any models that have been trained on this data set.

Comparison With Prior Work

Although multiple devices can estimate VO_2max, including several currently marketed consumer devices [35], the underlying data and algorithms are usually not published. The lack of data and method transparency limits the utility of these approaches for discovery-based research, where reproducibility is paramount. In contrast, an open approach to method validation can also serve as a foundation for downstream research in different conditions or populations to generate normative data for interpreting results [36].

As many dedicated hardware devices for digital health in the consumer sphere have experienced short half-lives of availability, we believe that the dependency only on a smartphone with a flash and camera may provide a greater degree of *future-proofing* for *heart snapshot*. This will be important for consistent, longitudinal measurements that may uncover patterns of VO_2max variance over time, especially in large-scale studies such as the AoURP.

Conclusions

The emerging development of consumer technology provides unprecedented opportunities to evaluate the use of additional digital biomarkers to improve risk management strategies for population health and for precision health at the level of an individual. Paired with access to large population studies, such as the AoURP [30] that collects health questionnaires, electronic health records, physical measurements, biospecimens, and digital health technology data, we can rapidly test emerging digital health measures for their potential to advance precision medicine. The *heart snapshot* software is freely available with all validation data and analysis code [37].

Acknowledgments

The authors would like to acknowledge Steve Steinhubl, Shannon Young, Nathaniel Brown, Joshua Liu, Erin Mounts, Stockard Simon, and Woody MacDuffie for their contributions to this work. Data are made available through Synapse (DOI: 10.7303/syn22107959).

Authors' Contributions

DEW and LO wrote the first draft of the paper. LO, MK, and JG developed the study and protocol, and MT developed algorithms for heart rate measurements. LO and MT performed the analyses. MH, DW, and JG recruited the participants and performed all measurements in the laboratory. MK and DEW oversaw the design and development of the heart snapshot apps. EA, VK, MVM, EM, JO, and LM helped identify the protocols for generalization, provided expert input, and edited the paper together with MK, LO, JG, DW, MT, MH, and DEW. LO and MK contributed equally to this study

Conflicts of Interest

EA is a founder and advisor for Personalis and Deep Cell and collaborates scientifically with Apple Inc. MVM is currently employed at Google.

Multimedia Appendix 1
Self-guided instructions and screen workflow for performing the heart snapshot VO₂max estimate.
[PNG File , 223 KB - mhealth_v9i6e26006_app1.png]

Multimedia Appendix 2

Demographic data for the maximal oxygen consumption validation study.
[XLSX File (Microsoft Excel File), 29 KB - mhealth_v9i6e26006_app2.xlsx]

References

1. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998 May 12;97(18):1837-1847. [doi: 10.1161/01.cir.97.18.1837] [Medline: 9603539]
2. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. J Am Med Assoc 2007 Feb 14;297(6):611-619. [doi: 10.1001/jama.297.6.611] [Medline: 17299196]
3. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. Br Med J 2017 May 23;357:j2099 [FREE Full text] [doi: 10.1136/bmj.j2099] [Medline: 28536104]
4. Laukkanen JA, Rauramaa R, Salonen JT, Kurl S. The predictive value of cardiorespiratory fitness combined with coronary risk evaluation and the risk of cardiovascular and all-cause death. J Intern Med 2007 Aug;262(2):263-272 [FREE Full text] [doi: 10.1111/j.1365-2796.2007.01807.x] [Medline: 17645594]
5. Mandsager K, Harb S, Cremer P, Phelan D, Nissen SE, Jaber W. Association of cardiorespiratory fitness with long-term mortality among adults undergoing exercise treadmill testing. JAMA Netw Open 2018 Oct 05;1(6):e183605 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.3605] [Medline: 30646252]
6. Ross R, Blair SN, Arena R, Church TS, Després J, Franklin BA, American Heart Association Physical Activity Committee of the Council on LifestyleCardiometabolic Health, Council on Clinical Cardiology, Council on EpidemiologyPrevention, Council on CardiovascularStroke Nursing, Council on Functional GenomicsTranslational Biology, Stroke Council. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. Circulation 2016 Dec 13;134(24):653-699. [doi: 10.1161/CIR.0000000000000461] [Medline: 27881567]
7. Guo Y, Bian J, Li Q, Leavitt T, Rosenberg EI, Buford TW, et al. A 3-minute test of cardiorespiratory fitness for use in primary care clinics. PLoS One 2018;13(7):e0201598 [FREE Full text] [doi: 10.1371/journal.pone.0201598] [Medline: 30059539]
8. Rippe JM, editor. Guidelines for exercise testing and prescription (ACSM). In: Encyclopedia of Lifestyle Medicine and Health. Thousand Oaks, California, United States: SAGE Publications Inc; 2012:1-1296.
9. Balady GJ, Arena R, Sietsema K, Myers J, Coke L, Fletcher GF, et al. Clinician's guide to cardiopulmonary exercise testing in adults. Circulation 2010 Jul 13;122(2):191-225 [FREE Full text] [doi: 10.1161/cir.0b013e3181e52e69]
10. McDavid A, Crane PK, Newton KM, Crosslin DR, McCormick W, Weston N, et al. Enhancing the power of genetic association studies through the use of silver standard cases derived from electronic medical records. PLoS One 2013;8(6):e63481 [FREE Full text] [doi: 10.1371/journal.pone.0063481] [Medline: 23762230]
11. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther 2000 Aug;80(8):782-807. [Medline: 10911416]
12. Bennett H, Parfitt G, Davison K, Eston R. Validity of submaximal step tests to estimate maximal oxygen uptake in healthy adults. Sports Med 2016 May;46(5):737-750. [doi: 10.1007/s40279-015-0445-1] [Medline: 26670455]
13. Muntaner-Mas A, Martinez-Nicolas A, Lavie CJ, Blair SN, Ross R, Arena R, et al. A systematic review of fitness apps and their potential clinical and sports utility for objective and remote assessment of cardiorespiratory fitness. Sports Med 2019 Apr;49(4):587-600 [FREE Full text] [doi: 10.1007/s40279-019-01084-y] [Medline: 30825094]
14. Cooper KH. A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. J Am Med Assoc 1968 Jan 15;203(3):201-204. [Medline: 5694044]
15. Hughes AD, Chaturvedi N. Estimation of maximal oxygen consumption and heart rate recovery using the tecumseh sub-maximal step test and their relationship to cardiovascular risk factors. Artery Res 2017 Jun;18:29-35 [FREE Full text] [doi: 10.1016/j.artres.2017.02.005] [Medline: 28546848]
16. Goodman JM, Thomas SG, Burr J. Evidence-based risk assessment and recommendations for exercise testing and physical activity clearance in apparently healthy individuals. Appl Physiol Nutr Metab 2011 Jul;36 Suppl 1:14-32. [doi: 10.1139/h11-048] [Medline: 21800940]
17. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
18. Montoye HJ, Block W, Keller JB, Willis PW. Fitness, fatness, and serum cholesterol: an epidemiological study of an entire community. Res Q Am Allian Health Phy Edu Recreat 2013 Mar 17;47(3):400-408 [FREE Full text] [doi: 10.1080/10671315.1976.10615390]
19. Milligan G. Fitness standards for the Maritime and Coastguard Agency and the oil and gas industry. University of Portsmouth. 2013. URL: https://pdfs.semanticscholar.org/739c/d165a84c7d7146a745580d0ad8e593f4dd3c.pdf [accessed 2021-05-13]
20. Eilers S, Bach DQ, Gaber R, Blatt H, Guevara Y, Nitsche K, et al. Accuracy of self-report in assessing Fitzpatrick skin phototypes I through VI. JAMA Dermatol 2013 Nov;149(11):1289-1294. [doi: 10.1001/jamadermatol.2013.6101] [Medline: 24048361]

21. Reeder AI, Hammond VA, Gray AR. Questionnaire items to assess skin color and erythemal sensitivity: reliability, validity, and "the dark shift". Cancer Epidemiol Biomarkers Prev 2010 May;19(5):1167-1173 [FREE Full text] [doi: 10.1158/1055-9965.EPI-09-1300] [Medline: 20447914]
22. Bino SD, Bernerd F. Variations in skin colour and the biological consequences of ultraviolet radiation exposure. Br J Dermatol 2013 Oct;169 Suppl 3:33-40. [doi: 10.1111/bjd.12529] [Medline: 24098899]
23. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb;327(8476):307-310. [doi: 10.1016/S0140-6736(86)90837-8]
24. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989 Mar;45(1):255. [doi: 10.2307/2532051]
25. CardiorespiratoryFitness-Android. Sage-Bionetworks. URL: https://github.com/Sage-Bionetworks/CardiorespiratoryFitness-Android [accessed 2020-11-24]
26. CardiorespiratoryFitness-iOS. Sage-Bionetworks. URL: https://github.com/Sage-Bionetworks/CardiorespiratoryFitness-iOS [accessed 2020-11-24]
27. Milligan G, House J, Tipton M, Hildenbrand B, Maeso M. A recommended fitness standard for the oil and gas industry. In: Proceedings of the European HSE Conference and Exhibition. 2013 Presented at: European HSE Conference and Exhibition; April 2013; London, United Kingdom. [doi: 10.2118/164960-MS]
28. Bandyopadhyay A. Validity of Cooper's 12-minute run test for estimation of maximum oxygen uptake in male university students. Biol Sport 2015 Mar;32(1):59-63 [FREE Full text] [doi: 10.5604/20831862.1127283] [Medline: 25729151]
29. Weisgerber M, Danduran M, Meurer J, Hartmann K, Berger S, Flores G. Evaluation of Cooper 12-minute walk/run test as a marker of cardiorespiratory fitness in young urban children with persistent asthma. Clin J Sport Med 2009 Jul;19(4):300-305. [doi: 10.1097/JSM.0b013e3181b2077a] [Medline: 19638824]
30. All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, et al. The "All of Us" research program. N Engl J Med 2019 Aug 15;381(7):668-676. [doi: 10.1056/NEJMsr1809937] [Medline: 31412182]
31. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017 May 24;7(2):3 [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
32. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18 [FREE Full text] [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
33. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 1988 Jun;124(6):869-871. [doi: 10.1001/archderm.124.6.869] [Medline: 3377516]
34. Ridder BD, Van Rompaey B, Kampen JK, Haine S, Dilles T. Smartphone apps using photoplethysmography for heart rate monitoring: meta-analysis. JMIR Cardio 2018 Feb 27;2(1):e4 [FREE Full text] [doi: 10.2196/cardio.8802] [Medline: 31758768]
35. Klepin K, Wing D, Higgins M, Nichols J, Godino JG. Validity of cardiorespiratory fitness measured with Fitbit compared to V˙O2max. Med Sci Sports Exerc 2019 Nov;51(11):2251-2256 [FREE Full text] [doi: 10.1249/MSS.0000000000002041] [Medline: 31107835]
36. Avram R, Tison GH, Aschbacher K, Kuhar P, Vittinghoff E, Butzner M, et al. Real-world heart rate norms in the Health eHeart study. NPJ Digit Med 2019;2:58 [FREE Full text] [doi: 10.1038/s41746-019-0134-9] [Medline: 31304404]
37. CRF validation analysis. Sage-Bionetworks. URL: https://github.com/Sage-Bionetworks/CRF_validation_analysis [accessed 2020-11-06]

Abbreviations

3-MST: 3-minute step test
12-MRT: 12-minute run test
ACF: autocorrelation function
AoURP: All of Us Research Program
EPARC: Exercise and Physical Activity Resource Center
UCSD: University of California, San Diego
VO_2max: maximal oxygen consumption

Edited by L Buis; submitted 24.11.20; peer-reviewed by M Altini, CP Lau; comments to author 11.01.21; revised version received 04.02.21; accepted 12.04.21; published 04.06.21.

Please cite as:
Webster DE, Tummalacherla M, Higgins M, Wing D, Ashley E, Kelly VE, McConnell MV, Muse ED, Olgin JE, Mangravite LM, Godino J, Kellen MR, Omberg L
Smartphone-Based VO2max Measurement With Heart Snapshot in Clinical and Real-world Settings With a Diverse Population: Validation Study
JMIR Mhealth Uhealth 2021;9(6):e26006
URL: https://mhealth.jmir.org/2021/6/e26006
doi:10.2196/26006
PMID:34085945

Corrigenda and Addenda

Correction: Tracking and Monitoring Mood Stability of Patients With Major Depressive Disorder by Machine Learning Models Using Passive Digital Data: Prospective Naturalistic Multicenter Study

Ran Bai[1,2*], MS; Le Xiao[3*], PhD, MD; Yu Guo[4], MEng; Xuequan Zhu[3], MA; Nanxi Li[3], MD; Yashen Wang[2], PhD; Qinqin Chen[2], PhD; Lei Feng[3], PhD, MD; Yinghua Wang[2], PhD; Xiangyi Yu[2], MS; Chunxue Wang[5], MD, PhD; Yongdong Hu[6], MD, PhD; Zhandong Liu[7], MD; Haiyong Xie[1,2], PhD; Gang Wang[1,3], PhD, MD

[1]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China

[2]National Engineering Laboratory for Risk Perception and Prevention, Beijing, China

[3]The National Clinical Research Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China

[4]Beijing University of Posts and Telecommunications, Beijing, China

[5]Department of Neuropsychiatry and Clinical Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

[6]Department of Psychological Medicine, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China

[7]Department of Neurology, Medical Health Center, Beijing Friendship Hospital, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Gang Wang, PhD, MD
The National Clinical Research Center for Mental Disorders
Beijing Anding Hospital
Capital Medical University
No 5 Ankang Lane, Xicheng District
Beijing, 100088
China
Phone: 86 13466604224
Fax: 86 1058303289
Email: gangwangdoc@ccmu.edu.cn

Related Article:

Correction of: https://mhealth.jmir.org/2021/3/e24365

(JMIR Mhealth Uhealth 2021;9(6):e30540) doi:10.2196/30540

In “Tracking and Monitoring Mood Stability of Patients With Major Depressive Disorder by Machine Learning Models Using Passive Digital Data: Prospective Naturalistic Multicenter Study” (JMIR Mhealth Uhealth 2021;9(3):e24365) the authors noted three errors.

1. In the originally published manuscript, there was no equal contribution footnote for authors Ran Bai and Le Xiao. This has been corrected to note that the authors contributed equally to the manuscript.

2. The affiliation for authors Le Xiao, Xuequan Zhu, Nanxi Li, Lei Feng, Gang Wang was incorrectly listed as follows in the original publication:

> *Beijing Anding Hospital, Capital Medical University, Beijing, China*

This has been corrected to the following:

> *The National Clinical Research Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China*

3. The original order of authors was as follows:

> *Ran Bai[1,2], MS; Le Xiao[3], PhD, MD; Yu Guo[4], MEng; Xuequan Zhu[3], MA; Nanxi Li[3], MD; Yashen Wang[2], PhD; Qinqin Chen[2], PhD; Lei Feng[3], PhD, MD; Yinghua Wang[2], PhD; Xiangyi Yu[2], MS; Haiyong Xie[1,2], PhD; Gang Wang[1,3], PhD, MD*
>
> *[1]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China*
>
> *[2]National Engineering Laboratory for Risk Perception and Prevention, Beijing, China*

[3]*Beijing Anding Hospital, Capital Medical University, Beijing, China*

[4]*Beijing University of Posts and Telecommunications, Beijing, China*

In the corrected version of the manuscript, Chunxue Wang, Yongdong Hu, and Zhandong Liu have been added as coauthors for their contribution to the design of protocol and data collection. All authors agree with the addition and new order of authors. The revised order of authors is as follows:

Ran Bai[1,2], MS; Le Xiao[3*], PhD, MD; Yu Guo[4], MEng; Xuequan Zhu[3], MA; Nanxi Li[3], MD; Yashen Wang[2], PhD; Qinqin Chen[2], PhD; Lei Feng[3], PhD, MD; Yinghua Wang[2], PhD; Xiangyi Yu[2], MS; Chunxue Wang[5], MD, PhD; Yongdong Hu[6], MD, PhD; Zhandong Liu[7], MD; Haiyong Xie[1,2], PhD; Gang Wang[1,3], PhD, MD*

[1]*Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China*

[2]*National Engineering Laboratory for Risk Perception and Prevention, Beijing, China*

[3]*The National Clinical Research Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China*

[4]*Beijing University of Posts and Telecommunications, Beijing, China*

[5]*Department of Neuropsychiatry and Clinical Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China*

[6]*Department of Psychological Medicine, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China*

[7]*Department of Neurology, Medical Health Center, Beijing Friendship Hospital, Capital Medical University, Beijing, China*

**these authors contributed equally*

The correction will appear in the online version of the paper on the JMIR Publications website on June 17, 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 19.05.21; this is a non–peer-reviewed article;accepted 19.05.21; published 17.06.21.

Please cite as:
Bai R, Xiao L, Guo Y, Zhu X, Li N, Wang Y, Chen Q, Feng L, Wang Y, Yu X, Wang C, Hu Y, Liu Z, Xie H, Wang G
Correction: Tracking and Monitoring Mood Stability of Patients With Major Depressive Disorder by Machine Learning Models Using Passive Digital Data: Prospective Naturalistic Multicenter Study
JMIR Mhealth Uhealth 2021;9(6):e30540
URL: https://mhealth.jmir.org/2021/6/e30540
doi:10.2196/30540
PMID:34138725

Original Paper

Contactless Sleep Monitoring for Early Detection of Health Deteriorations in Community-Dwelling Older Adults: Exploratory Study

Narayan Schütz[1], MSc; Hugo Saner[1,2,3], MD, PhD; Angela Botros[1], MSc; Bruno Pais[4], MSc; Valérie Santschi[4], PharmD, PhD; Philipp Buluschek[5], PhD; Daniel Gatica-Perez[6,7], PhD; Prabitha Urwyler[1,8], PhD; René M Müri[8], MD, PhD; Tobias Nef[1,8], PhD

[1]Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

[2]Department of Cardiology, University Hospital Bern, University of Bern, Bern, Switzerland

[3]I.M. Sechenov First Moscow State Medical University, Moscow, Russian Federation

[4]La Source, School of Nursing Sciences, HES-SO University of Applied Sciences and Arts of Western Switzerland, Lausanne, Switzerland

[5]DomoSafety SA, Lausanne, Switzerland

[6]Idiap Research Institute, Martigny, Switzerland

[7]École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[8]Department of Neurology, University Hospital Bern, University of Bern, Bern, Switzerland

Corresponding Author:
Tobias Nef, PhD
Gerontechnology and Rehabilitation Group
ARTORG Center for Biomedical Engineering Research
University of Bern
Murtenstrasse 50
Bern
Switzerland
Phone: 41 31 632 75 79
Email: tobias.nef@artorg.unibe.ch

Abstract

Background: Population aging is posing multiple social and economic challenges to society. One such challenge is the social and economic burden related to increased health care expenditure caused by early institutionalizations. The use of modern pervasive computing technology makes it possible to continuously monitor the health status of community-dwelling older adults at home. Early detection of health issues through these technologies may allow for reduced treatment costs and initiation of targeted preventive measures leading to better health outcomes. Sleep is a key factor when it comes to overall health and many health issues manifest themselves with associated sleep deteriorations. Sleep quality and sleep disorders such as sleep apnea syndrome have been extensively studied using various wearable devices at home or in the setting of sleep laboratories. However, little research has been conducted evaluating the potential of contactless and continuous sleep monitoring in detecting early signs of health problems in community-dwelling older adults.

Objective: In this work we aim to evaluate which contactlessly measurable sleep parameter is best suited to monitor perceived and actual health status changes in older adults.

Methods: We analyzed real-world longitudinal (up to 1 year) data from 37 community-dwelling older adults including more than 6000 nights of measured sleep. Sleep parameters were recorded by a pressure sensor placed beneath the mattress, and corresponding health status information was acquired through weekly questionnaires and reports by health care personnel. A total of 20 sleep parameters were analyzed, including common sleep metrics such as sleep efficiency, sleep onset delay, and sleep stages but also vital signs in the form of heart and breathing rate as well as movements in bed. Association with self-reported health, evaluated by EuroQol visual analog scale (EQ-VAS) ratings, were quantitatively evaluated using individual linear mixed-effects models. Translation to objective, real-world health incidents was investigated through manual retrospective case-by-case analysis.

Results: Using EQ-VAS rating based self-reported perceived health, we identified body movements in bed—measured by the number toss-and-turn events—as the most predictive sleep parameter (*t* score=–0.435, *P* value [adj]=<.001). Case-by-case analysis further substantiated this finding, showing that increases in number of body movements could often be explained by reported health incidents. Real world incidents included heart failure, hypertension, abdominal tumor, seasonal flu, gastrointestinal problems, and urinary tract infection.

Conclusions: Our results suggest that nightly body movements in bed could potentially be a highly relevant as well as easy to interpret and derive digital biomarker to monitor a wide range of health deteriorations in older adults. As such, it could help in detecting health deteriorations early on and provide timelier, more personalized, and precise treatment options.

(JMIR Mhealth Uhealth 2021;9(6):e24666) doi:10.2196/24666

KEYWORDS

sleep restlessness; telemonitoring; digital biomarkers; contactless sensing; pervasive computing; home-monitoring; older adults; toss and turns; sleep monitoring; body movements in bed

Introduction

In the face of an aging society in many countries, aging in place, often tied to home-care services, has attracted increasing public interest [1,2]. There is a clear preference and economic incentive for people to stay in their known environment [2,3]. Early institutionalization puts a significant additional burden on health care systems that are already struggling to keep up with the ongoing demographic changes [4]. An emerging trend in health care under these circumstances is the advent of pervasive computing technology [5-8]. Available research indicates good acceptance of such technology among older adults if the technology used does not require any interaction (ie, contactless) and is unobtrusive (ie, no voice or video recordings) [9-14]. This makes commonly used wearable devices, often encountered in younger demographics, suboptimal for long-term use with older adults.

Besides detecting acute events, such as falls, stroke, or acute heart failure, a potential application of this technology in older adults could be the detection of early signs of health issues [15]. Toward this goal, it is crucial to know which digital measures of health are useful early indicators of general health issues.

Compared with traditional biological biomarkers, objective health-relevant parameters obtained by means of modern information technology and outside of clinical environments are increasingly termed digital biomarkers [16-18]. Research in this domain is still relatively scarce but promising overall. Rantz et al [15], for instance, showed that using a variety of digital biomarkers they could successfully detect early changes in the health status of older adults. Moreover, they demonstrated that older adults with care enhanced through unobtrusive pervasive computing systems had better overall health outcomes compared with a control group. In the same line of research, Skubic et al [19] identified the most important features for the discrimination between useful and nonuseful health alerts as retrospectively judged by health care professionals. They found the most useful marker, in terms of useful to non-useful ratio, to be bed restlessness events.

Sleep is vital for normal physiologic functioning, with sleep problems becoming more prevalent with age. While some problems are thought to be a result of normal aging processes, many are indicative of underlying medical or psychological disorders or comorbidities [20]. In this context, Adam et al [21] found that community-dwelling older women with higher sleep onset delay (the time it takes to fall asleep) and lower sleep efficiency (fraction of time spent asleep while in bed) were significantly more likely to be placed in a nursing home. It has also been shown that sleep-related problems are an important factor driving institutionalizations among older adults and that they are often poorly recognized by caregivers [21,22]. As a result of people being stationary during sleep, contactless modern sensor technology can be used to measure sleep without device interactions or the necessity to wear and maintain devices likes smartwatches or fitness trackers [23-25]. In addition, sleep is much more comparable than daytime parameters as it generally occurs in the same setting, without many external factors being involved such as visitors, different activities, or different locations.

Considering these factors, previous findings, and the strong link between sleep and health in older adults, it seems that long-term sleep monitoring in older adults could be very useful for early detection and monitoring of health-related problems. However, further research investigating remote monitoring-derived sleep parameters as potential early indicators for general health issues is lacking. Toward this goal, we analyzed real-world longitudinal remote monitoring data from 37 older community-dwelling adults comprising more than 6000 nights worth of in-home sleep recordings and more than 600 matched health reports.

Methods

Overview

We performed a data-driven exploratory analysis of sleep data stemming from 37 older adult (aged 70 to 101 years) participants in Switzerland who were monitored for a target duration of at least 1 year, totaling 6686 recorded nights after preprocessing. Analysis aimed to identify sleep parameters that influenced health status. Since this is a hardly quantifiable and often subjective outcome, we used regularly reported perceived health reports as a proxy. To evaluate the effect on real-world health-related events, we further conducted manual case-by-case analysis, validating the most predictive sleep parameter of perceived health by using weekly health reports.

Participants

The analyzed data stems from 2 cohorts of Swiss seniors (pooled mean age 87 [SD 7] years; sex 67% [30/45] female) where novel computing technologies for aging in place scenarios were evaluated. Participants from both cohorts were recruited to represent naturalistic samples of community-dwelling, alone-living older adults in Switzerland [26,27]. Both studies took place between 2017 and 2018. The inclusion criteria were similar in that they aimed to recruit older adults aged 70 years and older who were living alone in an apartment or house without pets. In cohort 2, participants were additionally required to be followed by a local home-care association. While in cohort 1, only unwillingness to comply with the study protocol was an exclusion criterion, cohort 2 had the following exclusion criteria: (1) severe cognitive impairment making one unable to follow study protocol (clock-drawing score ≥4); (2) skin problems, such as irritations, itching, serious redness; (3) undergoing dialysis; (4) not willing to comply with the study protocol; (5) unable to understand the study aim; or (6) hospitalization planned in a short period of time [26]. The related studies were both conducted based on principles in the Declaration of Helsinki and approved by the Ethics Committee of the canton of Bern and Vaud, Switzerland (KEK-ID: 2016-00406 and CER-VD ID: 2016-00762, respectively). All subjects signed and returned an informed consent before study participation. We included all participants who started in either study, giving us a combined total of 45 participants (not all had matching data as some dropped out early).

Cohort Differences and Characteristics

In both studies, the same device for sleep recording and the same protocols for interviews and questionnaires were used. There are no differences between the study protocols with regard to the analyzed sleep parameters or health reports. The original studies were conducted by, in part, overlapping researchers but with distinct collaborators and local personnel (same concept, same subset of sensors, same questionnaires and reporting protocol, different technical personnel, different personnel conducting questionnaires and interviews, different personnel responsible for recruitment, different regions in Switzerland). Participant characteristics are shown in Table 1. Statistical significances of the respective baseline characteristics were evaluated using 2-sample, 2-sided *t* tests.

Table 1. Participant and questionnaire characteristics.

Characteristic	Cohort 1 (n=24)	Cohort 2 (n=21)	Cohort differences, *t* score (*P* value)	Pooled (n=45)
Age (years), mean SD	88 (7)	86 (7)	–0.76 (.50)	87 (7)
Sex, female, n (%)	19 (79)	11 (52)	1.94 (.06)	30 (67)
EQ-VAS[a], mean (SD)	78 (13)	72 (15)	1.35 (.19)	75 (14)
Nights measured, n (%)	4806 (87)	1880 (61)	—[b]	6686 (78)
Nights measured, mean (SD)	200 (104)	104 (97)	3.03 (.004)	159 (111)
Health reports, n (%)	963 (—)	803 (—)	—	1766 (—)
Health reports, mean (SD)	38 (10)	38 (24)	0.03 (.98)	38 (18)
Health reports, matched, n (%)	417 (43)	234 (29)	—	651 (37)
Health reports, matched, mean (SD)	18 (9)	15 (17)	0.56 (.58)	17 (13)

[a]EQ-VAS: EuroQol visual analog scale.

[b]Not applicable.

Ground Truth Data Acquisition

Perceived health and health-relevant events were recorded on a weekly basis by means of short interviews and questionnaires. These weekly reports were gathered as part of home visits and were mandatory for study participants to answer. To evaluate the participants' perception of health status, we used the well-validated EuroQol EQ-5D-3L health-related quality of life instrument.

The EQ-5D-3L is a well-established instrument to measure quality of life and has been shown to be consistently associated with mortality and hospitalization in older adults [28]. For simplicity reasons, we only used the EQ-VAS part of the EQ-5D-3L as it has been shown to exhibit stronger associations with number of hospitalizations and long-term mortality in older adults in comparison to the EQ-5D-3L index and classes of problems [28,29]. The EQ-VAS asks participants to describe their health state on a visual analog scale that ranges from 0 to 100 [30].

Sensor Data Acquisition

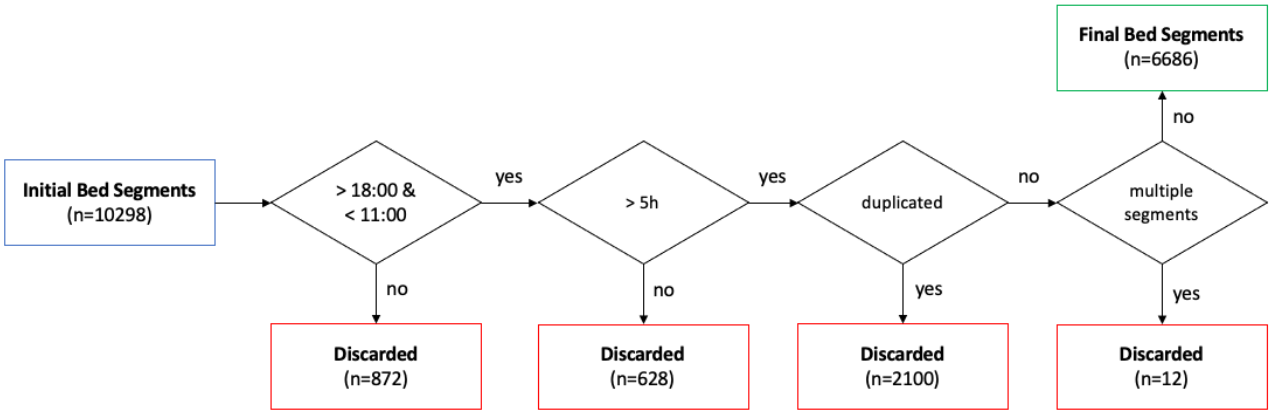
Sleep was monitored with a contactless, ferroelectret bed sensor (EMFIT QS, Emfit Ltd), which was placed under a person's mattress. The sensor comprises a thin quasi-piezoelectric film that translates thickness differentials into electrical charges [31]. The acquired sleep parameters were extracted by proprietary algorithms from the device manufacturer. Literature supports the accuracy of corresponding algorithms. It has notably been shown that the algorithms used with the device are capable of measuring heart rate and respiration rate with good accuracy [31]. Moreover, sleep staging algorithms of Emfit-based sensors have shown good agreement when compared with polysomnography-based ratings [32,33]. The devices were connected to the internet via specifically set up Wi-Fi hotspots.

Sleep data recorded at the participants' homes was sent to a cloud in real time and automatically acquired for analysis.

Data Processing

The acquired sleep segments were first filtered by their recording start and stop time (after 6 PM and before 11 AM) to exclude potential daytime naps from being included. Furthermore, only segments with a total duration of more than 5 hours were included, and duplicated segments were removed. Last, if multiple segments were recorded for a single night slot, the longest recording was used. Segments were defined as continuous bed activity with no more than 2 hours of out-of-bed time in between. A graphical overview is given in Figure 1.

Figure 1. Preprocessing flowchart.



Before modeling, all sleep data was standardized to zero mean and unit variance. To assess the relationship of sleep data with EQ-VAS ratings, sleep parameters of the night proceeding the questionnaire completion were matched with the respective EQ-VAS ratings. This resulted in a total of 651 matching data points across 37 participants. Every participant with at least one matching data point was included in the analysis.

Data Loss

Since the EMFIT sensors had to be connected to local Wi-Fi, numerous nights were lost due to connectivity issues with the mobile devices providing the hotspots (this includes loss of Wi-Fi connection, loss of 3G connectivity, network provider issues, and server issues). These problems were amplified by the logistical difficulty of fixing technical issues with remote monitoring in a timely manner. To reduce the risk of a potential bias as a result of missing nights, we evaluated the relationship between the percentage of missing matched data points and EQ-VAS ratings by means of the Spearman correlation coefficient of the 2 variables. To check for a potential bias introduced by the exclusion of participants without matching data points (primarily those who dropped out earlier), we performed a 2-sample, 2-sided *t* test to test for a significant difference in the means of EQ-VAS ratings between the participants included and the ones excluded as a result of not having matching data for this analysis.

Sleep Parameters

The analyzed sleep parameters include commonly measured quality-of-sleep metrics such as sleep efficiency, sleep latency, number of awakenings, rapid eye movement (REM) sleep, deep sleep, and light sleep [34]. Additionally, body movements, further referred to as the number of toss-and-turn events, as the parameter is named by the manufacturer, were analyzed. What counts as body movement is not exactly disclosed by the manufacturer but can be broadly described as a binary discretization (dichotomization) of the activity signal over short-time epochs of a few seconds. This is conceptually similar to what is being done when measuring activity counts with wearable accelerometers. A full overview of the included sleep parameters is given in Table 2 and Multimedia Appendix 1. All parameters are derived from proprietary algorithms from the manufacturer.

Table 2. Sleep parameters.

Parameter	Value, mean (SD)
Duration total (first to last recorded event in bed, sec)	31,823 (5776)
Duration in bed (sec)	30,788 (5607)
Average heart rate (beats per minute)	61.83 (6.32)
Average respiration rate (breaths per minute)	14.60 (2.78)
Number bed exits (count)	3.01 (2.50)
Number toss-and-turn events (larger movements, count)	57.62 (61.17)
Duration asleep (sec)	26,730 (5474)
Duration in REM[a] (sec)	6546 (2059)
Duration in light sleep (sec)	15,575 (3227)
Duration in deep sleep (sec)	4608 (1665)
Duration awake (sec)	4959 (2041)
Sleep onset delay (sec)	1446 (763)
Duration out of bed (sec)	1034 (914)
Heart rate variability high frequency band, 0.15-0.40 Hz (normalized power spectral density)	56.00 (11.35)
Heart rate variability low frequency band, 0.04-0.15 Hz (normalized power spectral density)	43.62 (11.13)
Number awakenings (count)	2.22 (1.44)
Percentage in deep sleep (%)	0.14 (0.04)
Percentage in REM sleep (%)	0.21 (0.05)
Percentage in light sleep (%)	0.49 (0.06)
Sleep efficiency (time asleep from total duration, %)	0.84 (0.07)

[a]REM: rapid eye movement.

Ranking of Relevant Sleep Parameters

To quantify the importance of individual sleep parameters on weekly reported EQ-VAS ratings, we used linear mixed-effects models (LMMs) with a random intercept per participant. For each sleep parameter, one model with EQ-VAS rating as response variable, corrected for age and sex, was fit. The rationale for using this type of model is based on the structure of the data: namely, multiple measures (with an uneven number of data points collected) per participant. Eventually, sleep parameters were ranked by the absolute value of their *t* scores. Reported *P* values were calculated on the basis of an implementation introduced by Kuznetsova et al [35]. We additionally report the respective *P* values when corrected for multiple comparisons, using the Bonferroni-Holm method. To provide insights into how the highest ranked predictor is associated with the remaining sleep parameters, individual Spearman correlation coefficients (*r*) were calculated. For all reported statistical hypothesis tests, we set a significance level of α=.05.

Case-by-Case Analysis of Toss-and-Turn Metric

While the above-mentioned ranking helps to identify the most relevant metric with regard to self-reported health, it does not necessarily imply a relevant effect on actual health. To check for an effect of the most relevant metric with regard to actual health, we performed manual, retrospective case-by-case analysis to investigate whether visually perceivable trends could be explained by health reports. For each manually identified trend (based on 14-day moving average) or point anomaly with regard to toss-and-turn time series, we analyzed the respective health reports around the same time to verify whether any health-relevant events that could explain the observations were reported.

Quantifying Anomalous Sleep Restlessness Evolutions

While manual inspection of the toss-and-turn metric may be possible in some settings, it is time-consuming and might thus not be practical. A relatively straightforward way to quantify structural changes in toss-and-turn counts, beyond just percentual changes, is to view them as the result of a Poisson process. In the case where the toss-and-turn number is approximately stable, this process can be seen as stationary (homogenous). In the presence of external factors, such as potential health changes, the toss-and-turn counts can be thought of as the result of a nonstationary Poisson process. To differentiate between the two, we assume the former and calculate the maximum likelihood estimate for the rate parameter λ_{TnT} to calculate the chi-square statistic χ^2 of the observed f_o and expected count frequencies f_e. In case the process is nonstationary, the chi-square statistic will become increasingly large and allows for a threshold to be set for automated notifications. To analyze the association of this approach with perceived health, we performed calculations based on weekly segmented toss-and-turn data and event counts in 30-minute

time intervals. EQ-VAS measures for the same week were matched and averaged.

Software Used

Data acquisition, preprocessing, and sleep parameter extraction were performed using the Python (Python Software Foundation) programming language. Statistical tests and LMMs were computed using the R statistical programming language (R Foundation for Statistical Computing) with package lme4 version 1.1-21 [36].

Results

Ranking of Relevant Sleep Parameters

We found the strongest predictor of perceived health to be the total number of toss-and-turn events per night (*t* score=–0.435, *P*<.001, *P* [adj]<.001). The only other predictor to remain significant after correcting for multiple comparison was the average nightly respiration rate (*t* score=–3.148, *P*=.002, *P* [adj]=.032). Both variables showed a negative association with EQ-VAS ratings. The full ranking is visualized in Figure 2, and all results are presented in Table 3. Overall associations between nightly toss-and-turn counts and the remaining sleep parameters based on Spearman correlation coefficients are shown in Figure 3. The strongest positive association was with the duration spent out of bed (*r*=.24) and the strongest negative one was found to be with the percentage spent in REM sleep (*r*=–.16). All associations were statistically significant, primarily as a result of the large number of 6686 data points.

Figure 2. Ranking of self-rated perceived health based on individual linear mixed effects models.

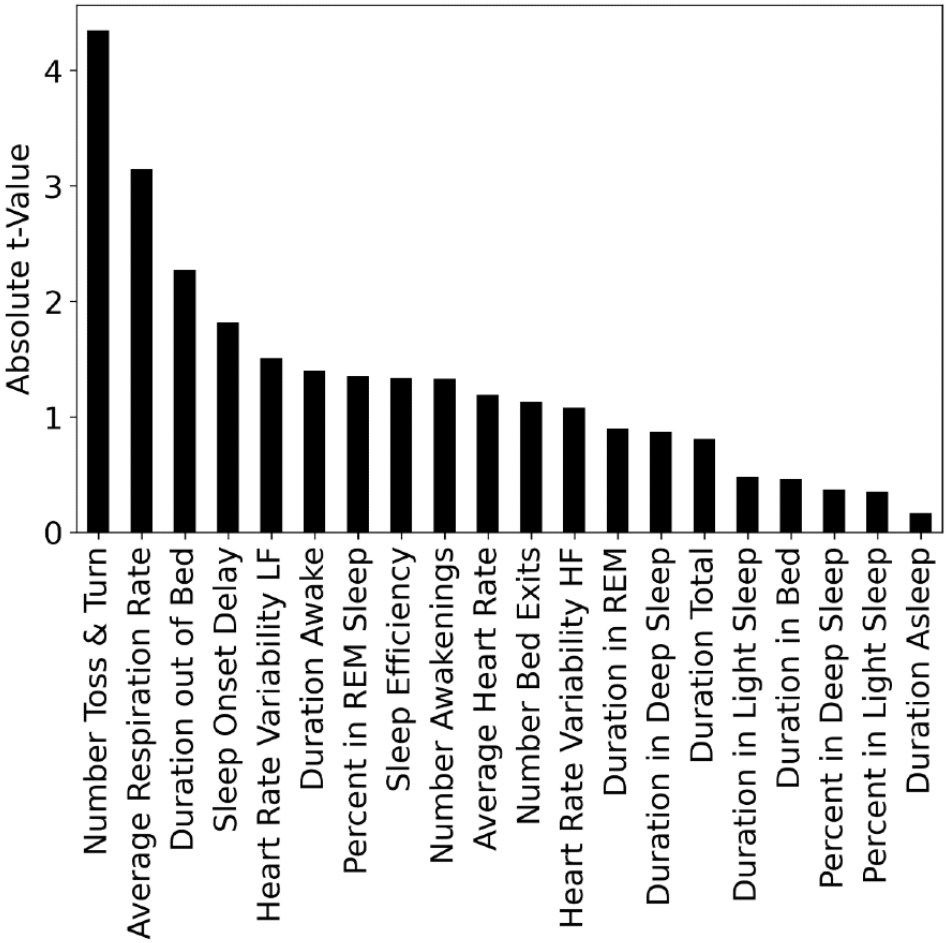
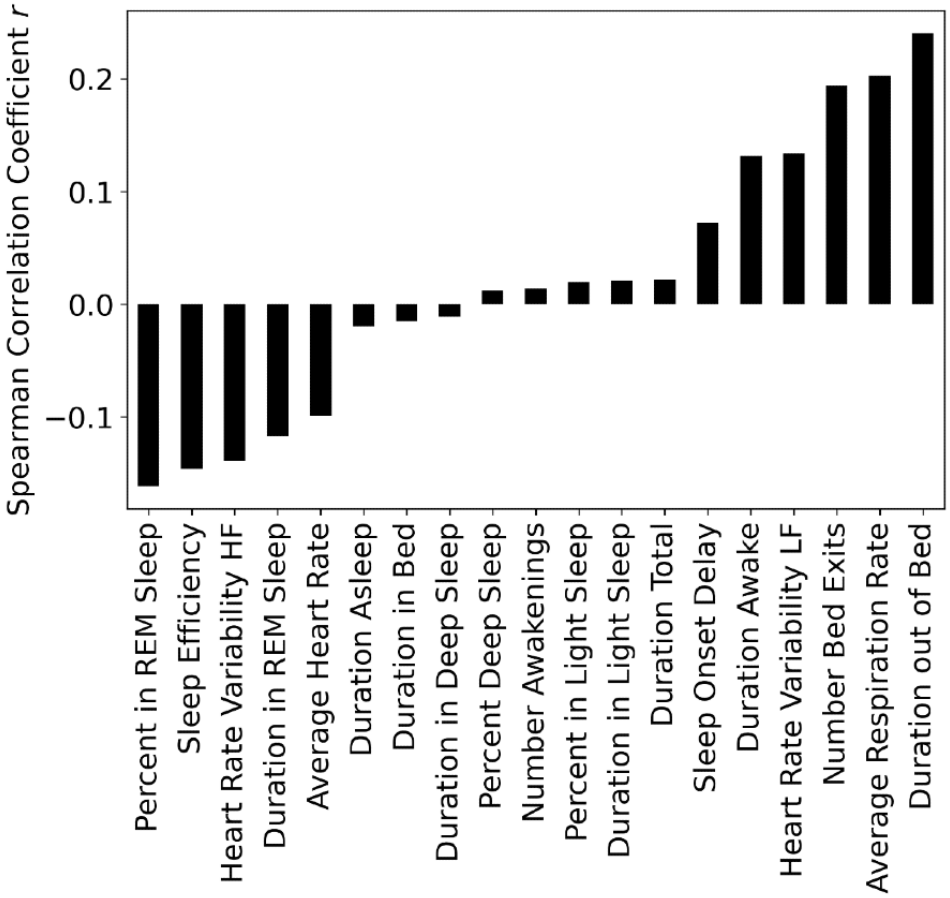


Table 3. Association of EuroQol visual analog scale ratings with sleep parameters based on mixed effects models.

Parameter	Estimate	*t* score	*P* value	*P* value (adjusted)
Number toss-and-turn events	–2.48	–4.35	<.001	<.001
Average respiration rate	–1.81	–3.15	.002	.03
Average heart rate	–0.73	–1.19	.23	>.99
Duration total	–0.38	–0.81	.42	>.99
Duration in bed	–0.21	–0.46	.65	>.99
Number bed exits	–0.55	–1.13	.26	>.99
Duration asleep	–0.08	–0.17	.86	>.99
Duration in REM[a]	0.36	0.90	.36	>.99
Duration in light sleep	–0.20	–0.48	.63	>.99
Duration in deep sleep	–0.30	–0.87	.39	>.99
Duration awake	–0.57	–1.40	.16	>.99
Sleep onset delay	–0.61	–1.82	.07	>.99
Duration out of bed	–1.03	–2.27	.02	.42
Heart rate variability high frequency band	–0.44	–1.08	.28	>.99
Heart rate variability low frequency band	0.61	1.51	.13	>.99
Number awakenings	0.57	1.33	.19	>.99
Percentage in deep sleep	–0.12	–0.37	.71	>.99
Percentage in REM sleep	0.49	1.35	.18	>.99
Percentage in light sleep	0.13	0.35	.73	>.99
Sleep efficiency	0.54	1.34	.18	>.99

[a]REM: rapid eye movement.

Figure 3. Relationship between number of toss-and-turns and the remaining sleep parameters.



Cohort Differences and Data Loss

Regarding potential biases in the above presented results, we found that the difference in participant characteristics of the analyzed cohorts was mostly insignificant (supporting the null hypothesis that the means of the 2 cohorts are equal) with the exception of the number of measured nights per person (*t* score=3.03, *P*=.004); more details can be found in Table 1. We further found a nonsignificant and very weak correlation between the percentage of missing nights and reported EQ-VAS ratings per participant (*r*=–.10, *P*=.54, n=37). Regarding participants who were not included in the analysis as a result of not having matching nights measured, we found no significant difference in the means of EQ-VAS ratings (*t* score=–0.44, *P*=.66).

Case-by-Case Analysis

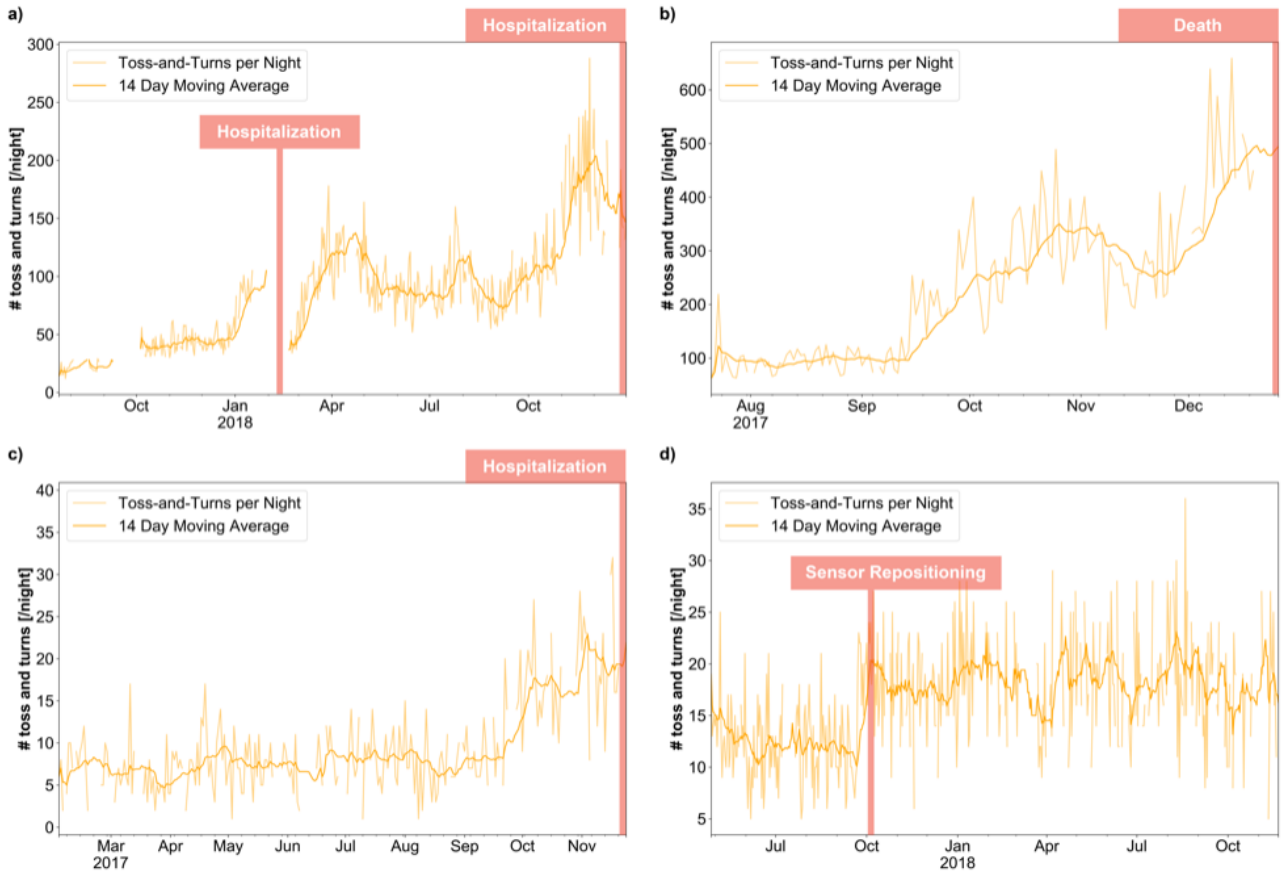
Results from qualitative case-by-case analysis, where trends and point anomalies with regard to the number of toss-and-turns per night were manually analyzed, are summarized in Table 4. The analysis showed that more than half (7/13, 54%) of abnormally looking toss-and-turn patterns could be related to reports of health-relevant events. Cases with a particularly high increase of more than 200 toss-and-turns per night (3 in total) were all accompanied by medically relevant events, of which 2 were severe and led to hospitalizations and death, respectively. Reported health incidents with visible change in toss-and-turns include heart failure, hypertension, abdominal tumor, seasonal flu, gastrointestinal problems, and urinary tract infection.

Four examples are displayed in Figure 4, where Figure 4A depicts the case of a participant who was hospitalized twice due to heart failure. Both hospitalizations were preceded by increases in the number of toss-and-turn events. It should be noted how the toss-and-turn numbers stabilized for a short time period after the first hospitalization and the resulting compensation. Figure 4B shows the number of nightly toss-and-turn events of a participant experiencing a rapid decline in health, later diagnosed as a large abdominal tumor and worsening heart failure, eventually leading to the participant's death. It is apparent how the number of toss-and-turn events per night increased strongly, with a small decrease prior to a final increase. Figure 4C shows the case of a participant who was hospitalized and subsequently institutionalized as a result of worsening hypertension problems. About 2 months prior to institutionalization, the number of toss-and-turn events started to increase markedly. Finally, in Figure 4D, a participant without any health issues is shown. The only visible (abrupt) signal change was due to a reported sensor repositioning.

Table 4. Summary of qualitative analysis of abnormal toss-and-turn patterns.

Identified Anomalous Patterns	Total cases, n	Cases with plausible explanation, n (%)
Trends (≤200 toss-and-turns)	13	7 (54)
Trends (>200 toss-and-turns)	3	3 (100)
Abnormal local peaks	20	13 (65)

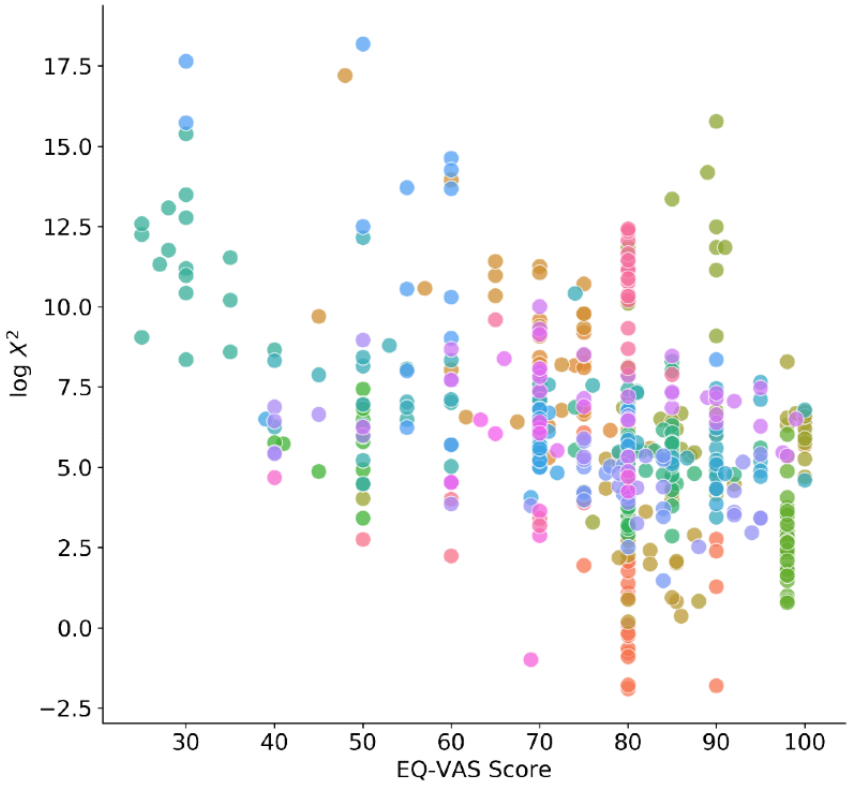
Figure 4. Evolutions of nightly toss-and-turn counts related to reported events.



Quantifying Anomalous Sleep Restlessness Evolutions

As can be seen in Figure 5, the log transformed chi-square statistics of weekly time intervals and matched EQ-VAS ratings show a negative association (r=–.42, P<.001, n=588) such that a higher chi-square indicates lower self-reported health. However, exact cutoff values that could be used to trigger alarms are difficult to find, since we do not have an absolute definition for relevant health status changes. We found that, in the analyzed population, weekly chi-square values greater than 1000 were mostly related to health changes and other sensor anomalies like displacements.

Figure 5. Relationship between chi-square statistics and EuroQol visual analog scale (EQ-VAS) ratings.



Discussion

Principal Findings

We retrospectively analyzed real-world, long-term remote monitoring sleep data from two studies in Switzerland where older adults' homes were equipped with modern pervasive computing systems and followed weekly by nursing staff. The goal of this explorative analysis was to find which unobtrusively measurable sleep parameter was most indicative of general health deteriorations in community-dwelling older adults. Among 20 daily extracted sleep parameters, the number of toss-and-turn events per night, which can be thought of as the amount of body movements, was found to exhibit the strongest association with weekly EQ-VAS based self-rated health ratings. It should be mentioned that this association is likely not only a subjective momentary feeling since the EQ-VAS is known to be strongly related to all-cause mortality and hospitalizations in older adults [28,29]. In accordance with this relationship, we found that the toss-and-turn parameter often exhibits very apparent changes that we could relate to reported health incidents. As such, all cases leading to more than 200 toss-and-turn events per night were related to reported health-relevant incidents. But smaller trends could also be explained by events mentioned in health reports in at least half of the cases. Encountered incidents include heart failure, hypertension, abdominal tumor, seasonal flu, gastrointestinal problems, and urinary tract infection.

Since long-term home-monitoring in older adults is a new field with very few research groups actually performing real-world studies, there is not much research to relate these findings to. However, one of the pioneering groups in this field reported a surprisingly similar finding. In 2015, Skubic et al [19] found sleep restlessness events measured by pneumatic transducers on top of the mattress to be the variable resulting in the highest ratio of good to poor health alerts in a population of older adults. They assessed a variety of parameters, including activity patterns in different rooms but also discretized heart and breathing rates during sleep. We want to point out that what they called bed restlessness events is conceptually the same as the body movements we call toss-and-turns. In the end, this is a question of nomenclature. While further investigation is certainly necessary, it is unlikely that the similarities in these two independent studies have occurred by chance. The exact reasons as to why body movements exhibit this association with a variety of health deteriorations remains to be elucidated. In theory, most of the involved health incidents are accompanied by inflammation. One potential mechanism could thus be related to immune system–modulated changes as a response to inflammatory processes [37]. Beyond inflammation, it is known that psychosocial stress can also have an impact on various sleep metrics like the number of awakenings or sleep efficiency [38]. Whatever the exact underlying causes, it is not a stretch to assume that increases in toss-and-turns come with more fragmented sleep, more awakenings, and more time spent awake in bed. This is also supported by the positive correlation of toss-and-turns with bed exits, time out of bed, number of awakenings or time spent awake. Additionally, it is known that older adults are at particular risk of being awoken by body movements during sleep – potentially further increasing the relevance of nightly toss-and-turns in this demographic [39].

An interesting aspect of these findings is that well-known sleep quality metrics like sleep duration and sleep efficiency are seemingly less relevant than body movements when it comes

to an overall association with perceived health in older adults. Correlation analysis between the toss-and-turn values and those metrics further shows weak overall correlations (see Figure 3) and thus suggests that toss-and-turns actually capture different components of sleep quality behavior.

When it comes to detecting relevant changes in toss-and-turns, interpreting the generation of toss-and-turn events as a Poisson process makes it straightforward to detect even smaller changes in a principled manner and allows for potentially triggering alarms at predefined threshold values. This provides a way of quantifying changes that is likely less prone to a specific implementation of an algorithm that discretizes movements in bed into individual (toss-and-turn) events.

The number of toss-and-turns per night is a sleep parameter that is fairly simple to measure unobtrusively, thus without any interactions or privacy-compromising measurement modalities. Nightly body movements in bed, as quantified by the number of toss-and-turns, could thus potentially be used as a convenient digital biomarker for early detection and monitoring of a wide range of health deteriorations in older adults, including potentially life-threatening conditions like heart failure or certain types of cancer. Early detection in many cases could help to involve medical professionals in early stages of disease and provide more timely treatment options and preventive measures [15,40]. In addition, having a responsive digital biomarker in such cases might also allow for monitoring of the effectiveness of treatments, indicating when a specific treatment plan is failing and needs adjustment. A real-world example of this is the case of heart failure as shown in Figure 4A, where decompensation in the hospital was showing initial relief but, as is well visible by the toss-and-turns, was only temporary and reclined shortly after hospital discharge. This deterioration necessitated further adjustment in medication. The need for such an adjustment might have been detectable earlier by monitoring markers like toss-and-turns proactively.

Given the relationship with EQ-VAS ratings, evidence from real health events, and the similar finding from Skubic et al [19], we find nightly body movements in older adults to be a very promising digital biomarker that needs further investigation. Ideally, future studies should compare against controlled polysomnography measurements and potentially also different devices. Finally, larger datasets in this regard could also allow for a statistically sound evaluation of the effect of sleep parameter combinations.

Limitations

Working in true remote-monitoring settings comes with downsides and very time-consuming logistics. As such, we had to deal with a nonnegligible number of missing nights, mainly due to connectivity issues with wireless networks. While this factor is most likely independent from participants and thus should not introduce a bias (as supported by the nonexistent association of EQ-VAS scores and missing nights) with respect to our results, it highlights practical problems with this specific setup used to collect the presented data. It should be noted that many of the issues we experienced can and will be solved eventually. Besides technical limitations, there are a few major shortcomings of the presented work.

First, while we did analyze a large number of nights, 37 participants is still a very limited sample size and we are thus not able to provide general information on which kind of health issues can be detected using the toss-and-turn metric.

Second, our analysis might be somewhat biased toward the proprietary algorithms of the device manufacturer. This should not invalidate findings but could mean that difficult-to-measure parameters like heart rate variability or exact sleep stages are exceedingly noisy and might in reality be more important than what we reported, especially as the reliability of such algorithms may deteriorate in presence of more movement.

Third, we did not have access to information about movement-related sleep disorders such as restless leg syndrome, which likely influence the toss-and-turn metric. It should, however, be noted that this would likely not affect the detection of relative changes, which we consider to be the most relevant.

Last, the pooling of the two studies might be a concern. However, apart from the number of missing nights, participant baseline characteristics showed no significant differences and the study protocols with regard to the results shown were the same. The difference in missing nights is explained by different hotspot technologies used and distinct technical personnel being responsible for maintenance. As such, this limitation should be considered but might not really have influenced our results.

Conclusion

In this study, we evaluated which contactlessly measurable sleep parameter exhibits the strongest association with EQ-VAS ratings in older adults and manually analyzed real-world health events with respect to the parameter with the strongest association. For this we used long-term sleep data acquired with contactless bed sensors in the homes of older community-dwelling adults. We found body movements in bed, quantified by the number of toss-and-turn events, to be the most predictive sleep parameter for EQ-VAS based perceived health ratings among 20 potential sleep parameters. Supporting this finding, increases in toss-and-turn events turned out to often be a precursor of reported real-world health incidents. Furthermore, these results are in accordance with an independent previous finding in literature. Monitoring body movements in bed could thus serve as an interesting and relatively easy to acquire and interpret digital biomarker, allowing health care professionals to proactively screen for and monitor early signs of numerous health deteriorations in older adults. While further evidence from larger, more targeted, studies will be necessary, the potential of such a digital biomarker to be used as digital care support measure might be significant and should be further investigated.

Acknowledgments

We would like to thank all subjects for their participation. In addition, we thank everyone involved in gathering the presented data. In particular, we would like to thank the Neuchâtel Organise le Maintien à Domicile (NOMAD) home care association for their invaluable help in data collection and Aileen Näf and Stephan Gerber for their valuable feedback. This study was partially funded by the Swiss Commission for Technology and Innovation through the SWISKO project (17662.2 PFES-ES).

Authors' Contributions

NS, HS, BP, VS, PB, DGP, PU, LM, RMM, and TN designed and planned the study. NS, HS, and BP installed and maintained the system and measured the participants. NS and AB analyzed the data. NS, AB, PU, and TN wrote the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

PB is employed by DomoSafety SA, which is a reseller of the described sensor system. The remaining authors declare no potential conflict of interest.

Multimedia Appendix 1
Boxplots of measured sleep parameters for all participants.
[PDF File (Adobe PDF File), 4513 KB - mhealth_v9i6e24666_app1.pdf]

References

1. Greenfield EA. Using ecological frameworks to advance a field of research, practice, and policy on aging-in-place initiatives. Gerontologist 2012 Feb;52(1):1-12. [doi: 10.1093/geront/gnr108] [Medline: 22075772]
2. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
3. Dupuis-Blanchard S, Gould O, Gibbons C, Simard M, Éthier S, Villalon L. Strategies for aging in place: the experience of language-minority seniors with loss of independence. Glob Qual Nurs Res 2015;2:1-9 [FREE Full text] [doi: 10.1177/2333393614565187] [Medline: 28462299]
4. de Almeida Mello J, Declercq A, Cès S, Van Durme T, Van Audenhove C, Macq J. Exploring home care interventions for frail older people in Belgium: a comparative effectiveness study. J Am Geriatr Soc 2016 Nov;64(11):2251-2256. [doi: 10.1111/jgs.14410] [Medline: 27676585]
5. Beard JR, Bloom DE. Towards a comprehensive public health response to population ageing. Lancet 2015 Feb 14;385(9968):658-661 [FREE Full text] [doi: 10.1016/S0140-6736(14)61461-6] [Medline: 25468151]
6. Martínez-Alcalá C, Pliego-Pastrana P, Rosales-Lagarde A, Lopez-Noguerola J, Molina-Trinidad E. Information and communication technologies in the care of the elderly: systematic review of applications aimed at patients with dementia and caregivers. JMIR Rehabil Assist Technol 2016 May 02;3(1):e5226 [FREE Full text] [doi: 10.2196/rehab.5226] [Medline: 28582258]
7. Jaschinski C, Ben Allouch S, Peters O, Cachucho R, van Dijk J. Acceptance of technologies for aging in place: a conceptual model. J Med Internet Res 2021 Mar 31;23(3):e22613 [FREE Full text] [doi: 10.2196/22613] [Medline: 33787505]
8. Au-Yeung W, Miller L, Beattie Z, Dodge H, Reynolds C, Vahia I, et al. Sensing a problem: proof of concept for characterizing and predicting agitation. Alzheimers Dement (N Y) 2020;6(1):e12079 [FREE Full text] [doi: 10.1002/trc2.12079] [Medline: 32864417]
9. Lentelink S, Tabak M, van Schooten B, Hofs D, op den Akker H, Hermens H. A remote home monitoring system to support informal caregivers of people with dementia. 2018 Presented at: 11th International Conference on Health Informatics; 2018; Funchal.
10. Mahmood A, Yamamoto T, Lee M, Steggell C. Perceptions and use of gerotechnology: implications for aging in place. J Housing Elderly 2008 Jun 16;22(1-2):104-126. [doi: 10.1080/02763890802097144]
11. Rantz MJ, Skubic M, Miller SJ, Galambos C, Alexander G, Keller J, et al. Sensor technology to support aging in place. J Am Med Dir Assoc 2013 Jun;14(6):386-391 [FREE Full text] [doi: 10.1016/j.jamda.2013.02.018] [Medline: 23562281]
12. Bujnowska-Fedak M, Grata-Borkowska U. Use of telemedicine-based care for the aging and elderly: promises and pitfalls. Smart Homec Technol Teleheath 2015 May;3:91. [doi: 10.2147/shtt.s59498]
13. Urwyler P, Stucki R, Rampa L, Müri R, Mosimann UP, Nef T. Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living. Sci Rep 2017 Dec 08;7:42084 [FREE Full text] [doi: 10.1038/srep42084] [Medline: 28176828]
14. Pais B, Buluschek P, DuPasquier G, Nef T, Schütz N, Saner H, et al. Evaluation of 1-year in-home monitoring technology by home-dwelling older adults, family caregivers, and nurses. Front Public Health 2020 Oct 2;8:518957 [FREE Full text] [doi: 10.3389/fpubh.2020.518957] [Medline: 33134236]

15. Rantz MJ, Skubic M, Popescu M, Galambos C, Koopman RJ, Alexander GL, et al. A new paradigm of technology-enabled vital signs for early detection of health change for older adults. Gerontology 2015 Nov 26;61(3):281-290 [FREE Full text] [doi: 10.1159/000366518] [Medline: 25428525]
16. Meister S, Deiters W, Becker S. Digital health and digital biomarkers: enabling value chains on health data. Curr Dir Biomed Eng 2016;2(1):581. [doi: 10.1515/cdbme-2016-0128]
17. Coravos A, Goldsack J, Karlin D, Nebeker C, Perakslis E, Zimmerman N, et al. Digital medicine: a primer on measurement. Digit Biomark 2019;3(2):31-71 [FREE Full text] [doi: 10.1159/000500413] [Medline: 32095767]
18. Coravos A, Khozin S, Mandl KD. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. NPJ Digit Med 2019;2(1):1 [FREE Full text] [doi: 10.1038/s41746-019-0090-4] [Medline: 30868107]
19. Skubic M, Guevara RD, Rantz M. Automated health alerts using in-home sensor data for embedded health assessment. IEEE J Transl Eng Health Med 2015;3:1-11. [doi: 10.1109/jtehm.2015.2421499]
20. Bloom H, Ahmed I, Alessi C, Ancoli-Israel S, Buysse D, Kryger M, et al. Evidence-based recommendations for the assessment and management of sleep disorders in older persons. J Am Geriatr Soc 2009 May;57(5):761-789 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02220.x] [Medline: 19484833]
21. Spira AP, Covinsky K, Rebok GW, Stone KL, Redline S, Yaffe K. Objectively measured sleep quality and nursing home placement in older women. J Am Geriatr Soc 2012 Jul 15;60(7):1237-1243 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04044.x] [Medline: 22702839]
22. Marasinghe C. Sleep disorders in elderly. J Cey Coll of Phy 2014 Apr 09;43(1-2):37. [doi: 10.4038/jccp.v43i1-2.6815]
23. Fallmann S, Chen L. Computational sleep behavior analysis: a survey. IEEE Access 2019;7:142421-142440. [doi: 10.1109/access.2019.2944801]
24. Aubert X, Brauers A. Estimation of vital signs in bed from a single unobtrusive mechanical sensor: algorithms and real-life evaluation. Annu Int Conf IEEE Eng Med Biol Soc 2008;2008:4744-4747. [doi: 10.1109/IEMBS.2008.4650273] [Medline: 19163776]
25. Perez-Pozuelo I, Zhai B, Palotti J, Mall R, Aupetit M, Garcia-Gomez J, et al. The future of sleep health: a data-driven revolution in sleep science and medicine. NPJ Digit Med 2020;3:42 [FREE Full text] [doi: 10.1038/s41746-020-0244-4] [Medline: 32219183]
26. Pais B, Buluschek P, DuPasquier G, Nef T, Schütz N, Saner H, et al. Evaluation of 1-year in-home monitoring technology by home-dwelling older adults, family caregivers, and nurses. Front Public Health 2020 Oct 2;8:518957 [FREE Full text] [doi: 10.3389/fpubh.2020.518957] [Medline: 33134236]
27. Saner H, Schütz N, Botros A, Urwyler P, Buluschek P, du Pasquier G, et al. Potential of ambient sensor systems for early detection of health problems in older adults. Front Cardiovasc Med 2020 Jul 15;7:110 [FREE Full text] [doi: 10.3389/fcvm.2020.00110] [Medline: 32760739]
28. Cavrini G, Broccoli S, Puccini A, Zoli M. EQ-5D as a predictor of mortality and hospitalization in elderly people. Qual Life Res 2012 Mar;21(2):269-280. [doi: 10.1007/s11136-011-9937-0] [Medline: 21656336]
29. Cavrini G, Zamberletti J, Zoli M. Could the EQ-5D be used to predict mortality and hospitalization over a long term period? Soc Indic Res 2015 Aug 2;128(2):813-834. [doi: 10.1007/s11205-015-1058-7]
30. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
31. Ranta J, Aittokoski T, Tenhunen M, Alasaukko-oja M. EMFIT QS heart rate and respiration rate validation. Biomed Phys Eng Express 2019 Jan 15;5(2):025016. [doi: 10.1088/2057-1976/aafbc8]
32. Kortelainen JM, Mendez MO, Bianchi AM, Matteucci M, Cerutti S. Sleep staging based on signals acquired through bed sensor. IEEE Trans Inform Technol Biomed 2010 May;14(3):776-785. [doi: 10.1109/titb.2010.2044797]
33. Kelly JM, Strecker RE, Bianchi MT. Recent developments in home sleep-monitoring devices. ISRN Neurol 2012 Oct 14;2012:1-10 [FREE Full text] [doi: 10.5402/2012/768794] [Medline: 23097718]
34. Ohayon M, Wickwire E, Hirshkowitz M, Albert S, Avidan A, Daly F, 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] [Medline: 28346153]
35. Kuznetsova A, Brockhoff P, Christensen R. Package: tests in linear mixed effects models. J Stat Soft 2017;82(13):1. [doi: 10.18637/jss.v082.i13]
36. Bates D, Mächler M, Zurich E, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015:e4. [doi: 10.1007/0-387-22747-4_4]
37. Imeri L, Opp MR. How (and why) the immune system makes us sleep. Nat Rev Neurosci 2009 Mar;10(3):199-210 [FREE Full text] [doi: 10.1038/nrn2576] [Medline: 19209176]
38. Kim E, Dimsdale JE. The effect of psychosocial stress on sleep: a review of polysomnographic evidence. Behav Sleep Med 2007;5(4):256-278 [FREE Full text] [doi: 10.1080/15402000701557383] [Medline: 17937582]
39. Giganti F, Ficca G, Gori S, Salzarulo P. Body movements during night sleep and their relationship with sleep stages are further modified in very old subjects. Brain Res Bull 2008 Jan 31;75(1):66-69. [doi: 10.1016/j.brainresbull.2007.07.022] [Medline: 18158097]

40. Lette M, Stoop A, Lemmens LC, Buist Y, Baan CA, de Bruin SR. Improving early detection initiatives: a qualitative study exploring perspectives of older people and professionals. BMC Geriatr 2017 Jun 23;17(1):132 [FREE Full text] [doi: 10.1186/s12877-017-0521-5] [Medline: 28645251]

Abbreviations

EQ-5D-3L: EuroQol health-related quality of life instrument
EQ-VAS: EuroQol visual analog scale
LMM: linear mixed-effects model
REM: rapid eye movement

Edited by L Buis; submitted 29.09.20; peer-reviewed by M van der Kamp, W Sun, R Smelror, L Castro; comments to author 16.11.20; revised version received 27.02.21; accepted 23.04.21; published 11.06.21.

Please cite as:
Schütz N, Saner H, Botros A, Pais B, Santschi V, Buluschek P, Gatica-Perez D, Urwyler P, Müri RM, Nef T
Contactless Sleep Monitoring for Early Detection of Health Deteriorations in Community-Dwelling Older Adults: Exploratory Study
JMIR Mhealth Uhealth 2021;9(6):e24666
URL: https://mhealth.jmir.org/2021/6/e24666
doi:10.2196/24666
PMID:34114966

Review

Considerations for the Design and Implementation of COVID-19 Contact Tracing Apps: Scoping Review

Esli Osmanlliu[1,2], MD; Edmond Rafie[3], BEng; Sylvain Bédard[4]; Jesseca Paquette[3], MSc; Genevieve Gore[5], MLiS; Marie-Pascale Pomey[3,4,6], MD, PhD

[1]Research Institute of the McGill University Health Centre, Montreal, QC, Canada

[2]Division of Pediatric Emergency Medicine, Department of Pediatrics, Montreal Children's Hospital, McGill University Health Centre, Montreal, QC, Canada

[3]Research Centre of the University of Montreal Hospital Centre, Montreal, QC, Canada

[4]Centre d'excellence sur le partenariat avec les patients et le public, Montreal, QC, Canada

[5]Schulich Library of Physical Sciences, Life Sciences, and Engineering, McGill University, Montreal, QC, Canada

[6]Department of Health Policy, Management and Evaluation, School of Public Health, University of Montreal, Montreal, QC, Canada

Corresponding Author:
Esli Osmanlliu, MD
Division of Pediatric Emergency Medicine, Department of Pediatrics
Montreal Children's Hospital
McGill University Health Centre
MUHC Glen Site
1001 Décarie Blvd
Montreal, QC, H4A 3J1
Canada
Phone: 1 514 412 4499
Fax: 1 514 412 4217
Email: esli.osmanlliu@mcgill.ca

Abstract

Background: Given the magnitude and speed of SARS-CoV-2 transmission, achieving timely and effective manual contact tracing has been a challenging task. Early in the pandemic, contact tracing apps generated substantial enthusiasm due to their potential for automating tracing and reducing transmission rates while enabling targeted confinement strategies. However, although surveys demonstrate public interest in using such apps, their actual uptake remains limited. Their social acceptability is challenged by issues around privacy, fairness, and effectiveness, among other concerns.

Objective: This study aims to examine the extent to which design and implementation considerations for contact tracing apps are detailed in the available literature, focusing on aspects related to participatory and responsible eHealth innovation, and synthesize recommendations that support the development of successful COVID-19 contact tracing apps and related eHealth technologies.

Methods: Searches were performed on five databases, and articles were selected based on eligibility criteria. Papers pertaining to the design, implementation, or acceptability of contact tracing apps were included. Articles published since 2019, written in English or French, and for which the full articles were available were considered eligible for analysis. To assess the scope of the knowledge found in the current literature, we used three complementary frameworks: (1) the Holistic Framework to Improve the Uptake and Impact of eHealth Technologies, (2) the Montreal model, and (3) the Responsible Innovation in Health Assessment Tool.

Results: A total of 63 articles qualified for the final analysis. Less than half of the selected articles cited the need for a participatory process (n=25, 40%), which nonetheless was the most frequently referenced item of the Framework to Improve the Uptake and Impact of eHealth Technologies. Regarding the Montreal model, stakeholder consultation was the most frequently described level of engagement in the development of contact tracing apps (n=24, 38%), while collaboration and partnership were cited the least (n=2, 3%). As for the Responsible Innovation in Health framework, all the articles (n=63, 100%) addressed population health, whereas only 2% (n=1) covered environmental considerations.

Conclusions: Most studies lacked fundamental aspects of eHealth development and implementation. Our results demonstrate that stakeholders of COVID-19 contact tracing apps lack important information to be able to critically appraise this eHealth innovation. This may have contributed to the modest uptake of contact tracing apps worldwide. We make evidence-informed recommendations regarding data management, communication, stakeholder engagement, user experience, and implementation strategies for the successful and responsible development of contact tracing apps.

(JMIR Mhealth Uhealth 2021;9(6):e27102) doi:10.2196/27102

KEYWORDS

COVID-19; contact tracing; exposure notification; app; design; implementation; participatory; eHealth; surveillance; monitoring; review

Introduction

Background

As the global battle against the COVID-19 pandemic continues, the SARS-CoV-2 virus has infected over 160 million people and claimed over 3.3 million lives by May 2021 [1]. Even as vaccination programs expand their reach, nonpharmaceutical interventions such as social distancing, isolation, and quarantining remain essential for reducing viral transmission. Many will not have access to the vaccine until later in 2021, as limited production and delivery capacity warrants prioritizing high-risk groups. Many countries risk even longer delays in gaining access, despite efforts to promote a more equitable global distribution of the vaccine [2]. Moreover, long-term immunity from vaccination is not assured, and new virus variants have been associated with increased contagiosity [3]. Nonpharmaceutical interventions must therefore continue to support the fight against COVID-19, in complement to the global ramping up of vaccination.

Contact tracing is a fundamental containment strategy in response to emerging outbreaks. Public health agencies aim to rapidly identify individuals who may have been exposed to a person who is infected to recommend the most appropriate course of action (eg, self-isolation, symptom recording, and testing). The incubation period of this virus can last up to 14 days, during which infected individuals can unsuspectingly contaminate others [4,5]. Presymptomatic transmission along with other epidemiological, social, economic, and political challenges [6], coupled with chronic underfunding of public health systems [7], have undermined the reach of manual contact tracing during the COVID-19 pandemic.

Contact Tracing Apps as eHealth Solutions During the COVID-19 Pandemic

A variety of eHealth solutions, which leverage information and communication technologies for the betterment of health and health care services [8], have been proposed in response to the pandemic [9]. Early in the pandemic, digital contact tracing rapidly emerged as a promising tool to support manual tracers [10,11] and enable a more selective approach to regional lockdowns [12]. Contact tracing apps constitute an example of eHealth aimed at supporting standard nonpharmaceutical interventions. Generally, the intended use is to digitally collect information within their network of users to reduce pathogen transmission. A recent review of international technological innovations developed in response to the pandemic listed almost 100 tracing applications at different stages of development, most of which were smartphone-based [13]. These can be further categorized as position tracking applications, which aim to enforce the quarantine of infected individuals, and the more commonly deployed contact tracing applications, which continuously measure distances between users to rapidly notify the high-risk contacts of an individual with a confirmed SARS-CoV-2 infection. The latter type predominates internationally, as it is more compatible with civil liberties. Although the term “contact tracing app” is typically used in academic texts and the general press, the primary purpose of such tools is to rapidly identify and notify individuals who have had a high-risk exposure [10]. Terminology matters as it may erode public trust in tools that are perceived to enable state surveillance of individual mobility (*tracing*, *tracking*). In this context, the contact tracing technology codeveloped by Apple and Google refers to an “Exposure Notification System” [14]. Similarly, the government of Canada encourages the public to use COVID Alert, an “exposure notification app,” which it describes as “an additional tool to protect yourself and your loved ones” [15].

Contact tracing apps can differ according to eight fundamental characteristics. First, their installation can be voluntary or compulsory. Second, the extent of informed consent varies between apps. Third, some apps use a decentralized data management strategy, while others enable linkages with governmental agencies. Fourth, their ability to detect contact between users can rely on technologies such as GPS, Bluetooth, or Quick Response codes. Fifth, the specific algorithms deployed in the back end of these apps will determine their output (eg, the calculation of a risk of infection or the tracing of potential contacts). Sixth, they require varying levels of human oversight, if at all present. Seventh, the degree of interaction with users regarding recommended actions (eg, testing and isolation recommendations) and the extent of interaction with public health agencies can differ. Last, safety protocols for data privacy may also vary [13,16,17]. Contact tracing apps thus refer to a heterogeneous cluster of eHealth tools that likely differ in effectiveness and uptake depending on their respective design characteristics.

App Effectiveness, Barriers to Adoption, and Facilitators

Given their fundamental mechanism, the effectiveness of contact tracing apps depends in part on the level of uptake and ongoing use [10,18] by patients who have contracted the SARS-CoV-2 virus and other citizens. An influential modeling study published

by a team of researchers at the University of Oxford originally suggested that a 60% adoption rate should be targeted for effective virus transmission reduction, although any level of uptake may help lower disease transmission [19,20]. Even though numerous public consultations have suggested a general willingness to use contact tracing apps during the pandemic [21-28], the available data suggests low rates of continuous use in practice [27,29,30]. For instance, app penetration rates as of March 2021 were as low as 3.6% in France, 6.1% in Japan, 14% in Canada, and 28.5% in the United Kingdom. On the other hand, other countries such as Iceland and Finland have seen higher rates of adoption, at 38.5% and 45.3%, respectively [31,32]. The significant disparity in adoption rates worldwide highlights the fact that certain approaches could be better than others for successful implementation of this emerging technology.

As contact tracing apps encompass many underlying principles and disciplines, multiple aspects can facilitate or hinder their adoption. One of the main caveats in their implementation is concerns over data security and management. Societies are rightfully preoccupied with the challenges in reconciling civil liberties with public health imperatives in a pandemic context [33]. Data privacy, breaches in confidentiality, and the fear of mass surveillance are among the main concerns raised in surveys on user perspectives [21-24,27,28]. Moreover, as with other health informatics interventions, contact tracing apps may generate or exacerbate inequalities if they are not deployed carefully [34,35]. There is a risk of discrimination, repression, and systematic exclusion, especially among communities of color and marginalized groups, which are disproportionately affected by COVID-19 due to structural economic, political, and social vulnerabilities [35-37].

Key factors have been suggested as drivers for widespread success of contact tracing apps: integration with local health policy, adaptable workflows in an ever-evolving context, rapid notification systems, the ability to evaluate the effectiveness of the app transparently, and clear communications addressing privacy concerns [38].

New eHealth initiatives have emerged at an accelerating pace in the last decade; some have seen widespread adoption, whereas others have failed to provide sustained value. These failings can be attributed to design and implementation efforts that were initiated without a good understanding of the interdependencies between technology, societal values, and user experience in a health care setting. Many conceptual frameworks based on implementation science have been developed to evaluate and orient eHealth delivery. These frameworks highlight key factors that predict successful and sustainable eHealth technologies. The urgency of the ongoing public health crisis stimulated the rapid development of contact tracing apps and other eHealth innovations, and this generated a substantial number of related publications. Their coverage of essential design and implementation characteristics for eHealth innovation remains underinvestigated.

Objectives and Research Questions

The primary objective of this review is to map and analyze the literature on the design and implementation of COVID-19 contact tracing apps. This was achieved through three distinct questions: (1) to what extent does the available literature discuss features that promote the use of contact tracing apps by interested parties? (2) how have patients and citizens been engaged in the design and implementation of these apps? and (3) does the development of these apps correspond to principles of responsible research and innovation?

Through these questions, we studied how the development of contact tracing apps has taken into account considerations related to the uptake and impact of the innovation, the engagement of end users, and the responsible development of eHealth technologies. We ultimately identified the components required for successful and responsible eHealth development as described in the available literature on contact tracing apps and those that are lacking.

Methods

This study is reported according to PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [39] (Multimedia Appendix 1). The scoping review was conducted in accordance with the multistage framework outlined by Arksey and O'Malley [40], as detailed in the following sections.

Theoretical Frameworks

Three complementary frameworks were selected to address each of the research questions. To answer the first research question on factors that promote uptake of contact tracing apps, we selected the Framework to Improve the Uptake and Impact of eHealth Technologies proposed by van Gemert-Pijnen et al [41]. This framework advocates for a holistic approach comprising six guiding principles and requirements for successful eHealth technology development: (1) participatory processes (eg, upstream involvement of citizens in the selection of app features), (2) continuous evaluation cycles (eg, evaluation of the use and effect of the app on an ongoing basis to update risk assessments), (3) specific actions for implementation (eg, postdesign activities to promote or maintain app uptake), (4) foresight of changes in the organization of health care (eg, interaction of contact tracing apps with traditional public health processes), (5) persuasive design techniques (eg, technology-based suggestions to stimulate the uptake of contact tracing apps), and (6) advanced methods to assess impact (eg, key performance indicators to assess the impact of contact tracing apps). Given the novel application of contact tracing apps in the context of the COVID-19 pandemic, information regarding the latter stages of the framework may understandably be theoretical or perhaps even lacking.

For the second research question, which addresses the engagement of users in the development of contact tracing apps, we selected the Montreal model established by Pomey et al [42], as it focuses on the experiences and knowledge of patients and citizens regarding their health, trajectory of care, and related services [43]. It draws on the Patient and Family Engagement Framework proposed by Carman et al [44] to include a multidimensional assessment of patient engagement in health innovation [45]. For this study, we defined patients as

individuals who contracted the SARS-CoV-2 virus. This framework proposes four distinct levels along a patient or citizen engagement continuum, in ascending order of engagement: (1) information (eg, end users have the option of viewing an explanatory video on the proposed app), (2) consultation (eg, a survey to evaluate propensity for app acceptance), (3) collaboration (eg, app adaptation based on recommendations from the public), and (4) coconstruction (eg, governments engage stakeholders to co-design the app). Selected studies that referred to a participatory process were therefore further characterized according to the level of engagement they described.

Finally, for the third research question relating to responsible health innovation, we selected the Responsible Innovation in Health Assessment Tool by Pacifico Silva et al [46]. This framework comes from the field of Responsible Research and Innovation and helps assess health innovations by addressing challenges such as sustainability and equity. The framework elucidates five value domains that need to be considered for responsible innovation: (1) the population health domain, which includes the subvalues of health relevance, ethical, legal and social issues, and health equity (eg, how the app can be made accessible to vulnerable groups); (2) the health system value domain, which includes subsections on inclusiveness, level of care, and responsiveness (eg, contact tracing apps reduce labor from manual contact tracing); (3) the economic domain (eg, cost-effectiveness of contact tracing apps as compared to other public health interventions); (4) the organizational domain (eg, business strategies to increase app value); and (5) the environmental domain (eg, how app architecture can help reduce the carbon footprint).

Searching for Relevant Studies

A systematic literature search was performed by one author (GG) in five databases (PubMed, Scopus, IEEE Xplore, AMC Digital Library, and Europe PMC) using terms specifically related to the research question. Textbox 1 describes the search strategy used on PubMed. Additional database search strategies are available in Multimedia Appendix 2. The search was first performed on August 26, 2020, and produced 829 results. A second iteration performed on November 6, 2020, generated 1130 results, for a total of 1959 articles.

Textbox 1. Search strategy used on PubMed ([mesh] stands for Medical Subject Headings and indicates that the subject is indexed in the literature; [tw] indicates searches in the title and abstract fields).

((“contact tracing” [mesh] OR “epidemiological monitoring” [mesh]) AND (“mobile applications” [mesh] OR “algorithms” [mesh] OR “computer security” [mesh] OR “big data” [mesh] OR “computer simulation” [mesh] OR “geographic mapping” [mesh] OR “geographic information systems” [mesh] OR “microcomputers” [mesh] OR “software” [mesh]))

OR (tracing [tw] AND (app [tw] OR apps [tw] OR proximity [tw]))

OR ((contact [tw] OR exposure [tw]) AND notification* [tw] AND (app [tw] OR apps [tw] OR application* [tw]))

OR ((digital* [tw] OR mobile [tw] OR ehealth [tw] OR “eHealth” [tw] OR mhealth [tw] OR “m-health” [tw] OR app [tw] OR apps [tw] OR application* [tw] OR “geolocation*” [tw] OR “location service*” [tw] OR “location system*” [tw] OR “location information” [tw] OR gps [tw] OR big data [tw] OR ((geographic [tw] OR geographical [tw]) AND tracking [tw])) AND (“contact tracing” [tw] OR “contact tracking” [tw] OR “digital epidemiology” [tw]))

Selecting Studies

The study eligibility criteria (Textbox 2) were informed by a priori knowledge and by the review process itself, in keeping with the scoping review methodology. After removing duplicates, the inclusion and exclusion criteria were applied to the remaining articles.

Each of the 1959 articles was screened by one of two authors (ER and JP), initially excluding those with titles and abstracts unrelated to the topic of study. They then read the remaining articles to determine their relevance to the research question with respect to the inclusion criteria. The senior authors (EO and MPP) screened articles with uncertain relevance for final inclusion or exclusion. Discrepant decisions were resolved through team discussion and consensus. This process resulted in the inclusion of 63 articles, following the search iterations previously described.

Textbox 2. Eligibility criteria.

Inclusion criteria

- Related to design considerations of contact tracing apps for COVID-19 (eg, privacy considerations, citizen inclusion and participatory approach, or incentivization)
- Related to user experience or implementation approaches of contact tracing apps for COVID-19 (eg, surveys or focus groups, strategies to mitigate social vulnerabilities, or recommendations for governments communication with the public)

Exclusion criteria

- Published before 2019
- Published in a language other than French or English
- Pertaining exclusively to aspects of computer science or technical developments of contact tracing apps
- Pertaining exclusively to epidemiological feasibility or efficacy of contact tracing apps
- Authors could not obtain access to the full article

Charting the Data

Data from the included studies was charted on an extraction grid (Multimedia Appendix 3) according to the following categories: authors, title, date of publication, publication stage, aims of the study, and key findings. Articles were also classified into six categories: (1) proposal; (2) comment, editorial, or opinion piece; (3) survey or focus group; (4) case study; (5) review; and (6) essay. A description of each article type is provided in Textbox 3.

Textbox 3. Description of the types of articles.

Proposal

Considers the problems of a particular situation and offers a corresponding solution (eg, proposal to incentivize a contact tracing app [47])

Comment, editorial, or opinion piece

Reflects the author's or journal's opinion about a subject (eg, contact tracing app effectiveness and data security [48])

Survey or focus group

Concentrates on survey or focus group methods for data collection (eg, user acceptability of a contact tracing app [49])

Case study

Studies a particular *case* in depth (eg, development of the Trace Together app [50])

Review

Examines what has already been discovered about a subject (eg, systematic evaluation of content and features of a contact tracing app [51])

Essay

Discusses ideas from the literature in a support of arguments about a specific subject (eg, discussion of an intervention to introduce contact tracing technology [52])

Articles that outlined theoretical recommendations or criticism regarding the overarching concept of contact tracing apps were categorized as being based on theory (eg, ethical considerations of instantaneous contact tracing). In contrast, articles that factually described specific use cases or empirical studies were categorized as being based on practice (eg, description of technical features of a given app and their potential impact on implementation or surveys on user acceptability of a contact tracing app).

Furthermore, the extraction grid included the components of the Gemert-Pijnen et al [41], Pomey et al [42], and Pacifico Silva et al [46] frameworks, as previously discussed. One author (JP) coded each component to determine whether a given article addressed a framework (0 if not present, 1 if present) and extracted the supporting sentences, where applicable. A second author (ER) reviewed each component attribution and supporting sentences, as well as the theory or practice categorization, and reviewed the text for missing information. The principal authors then reviewed the analysis before reporting the results.

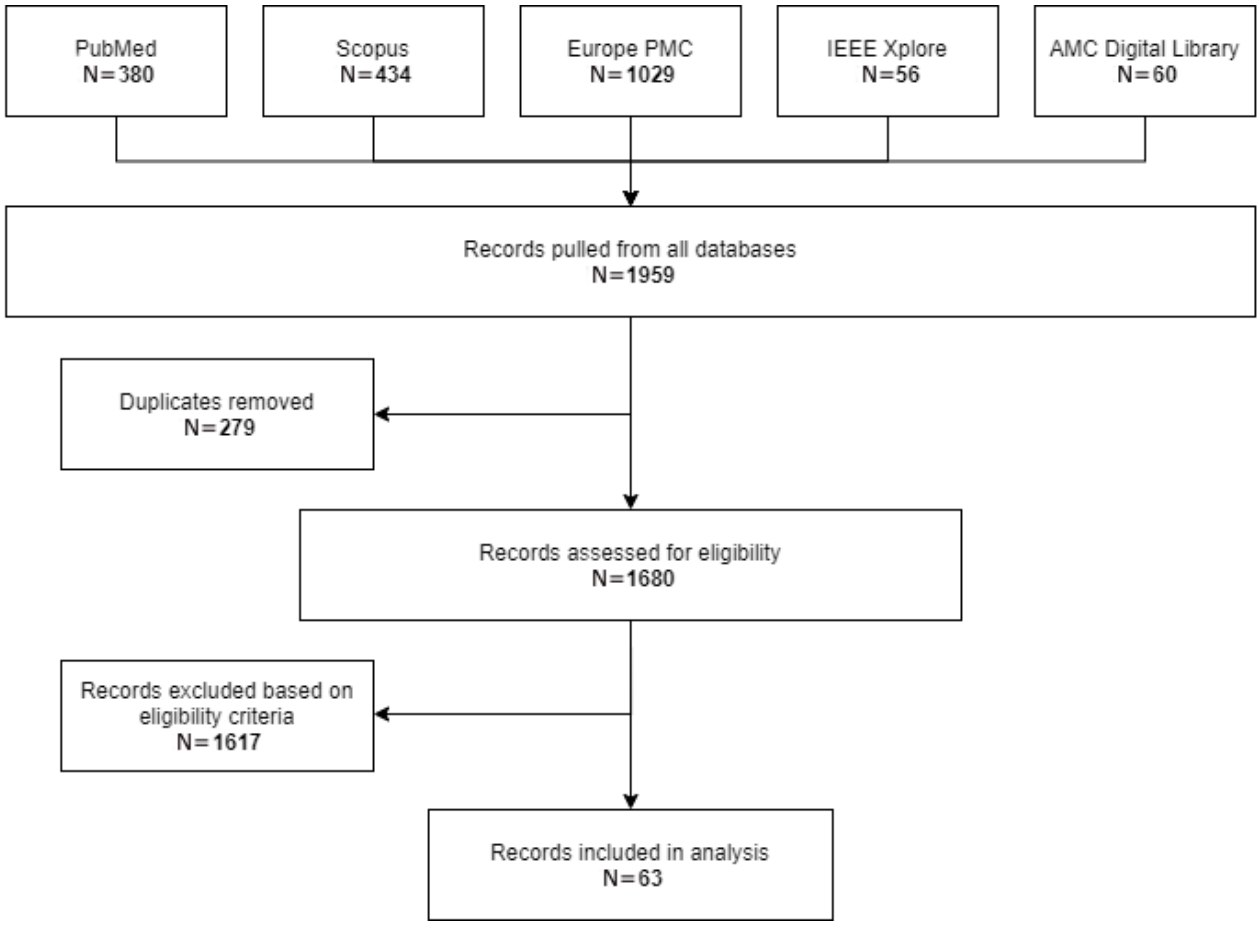
Collating, Summarizing, and Reporting the Results

Figures were produced using codes in accordance with the three theoretical frameworks used and other relevant information such as the date of publication, the type of article, and the theory or practice classification. Visualization of these elements shed light on changes in the available information and on gaps in the underreported domains of design and implementation of contact tracing apps. Critical appraisals of the included articles were beyond the scope of this study.

Results

This scoping review generated a total of 1959 records. Following the removal of duplicates and the application of the selection criteria, 63 articles were included in the analysis (Figure 1).

Figure 1. Flowchart of the study selection process.



Characteristics of Included Studies

The included studies (N=63) were published between April 16 and November 6, 2020, in 40 different journals or preprint databases. The majority (n=48, 76%) of included studies were published articles or e-prints, and 24% (n=15) were preprints. The number of published articles peaked in August (n=16, 25%; Figure 2). Theory-based studies were predominant from April to June, whereas the proportion of practice-based articles increased thereafter.

Figure 3 illustrates the types of articles by month of publication. Among the 63 included studies, most were surveys (n=16, 25%), followed by proposals and opinion pieces, editorials, or commentaries (n=13, 21%); reviews (n=10, 16%); essays (n=8, 13%); and case studies (n=3, 5%). Opinion pieces, editorials, and commentaries (in blue) were the predominant category before July 2020, followed thereafter by more diversity in publication types. Only surveys (n=1) were illustrated in our analysis in November, but this could be explained by the fact that our second and last iteration was carried out on November 6, 2020.

Figure 2. Number of monthly publications/preprints, according to theory- and practice-based categorization.

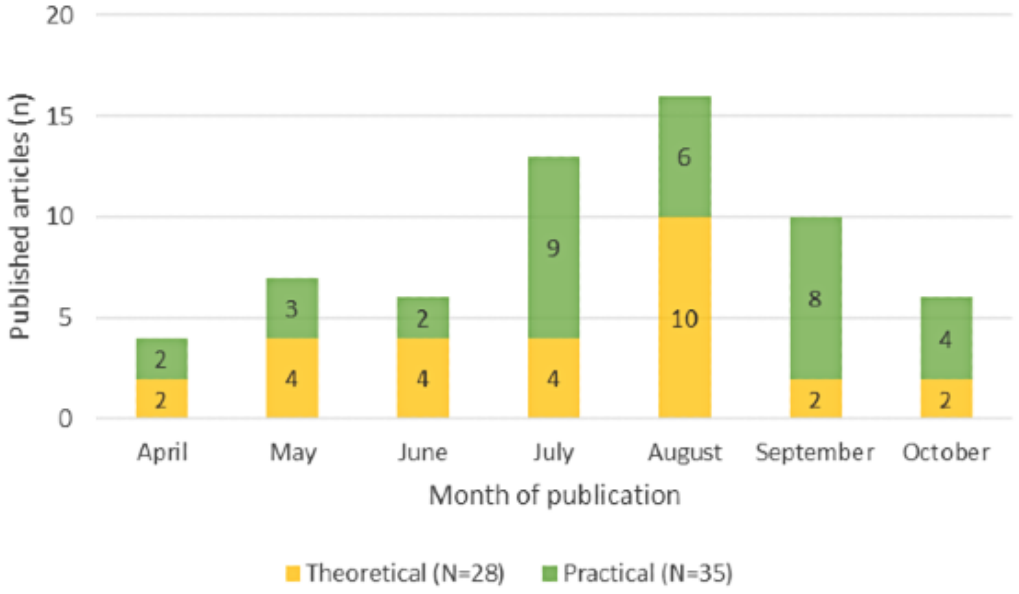
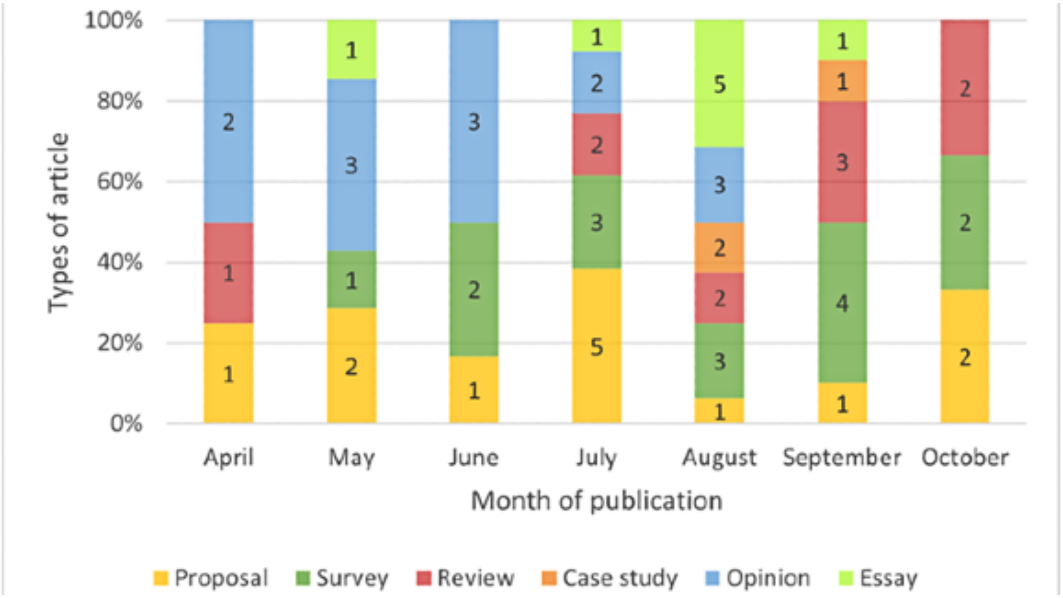


Figure 3. Types of article by month of publication.

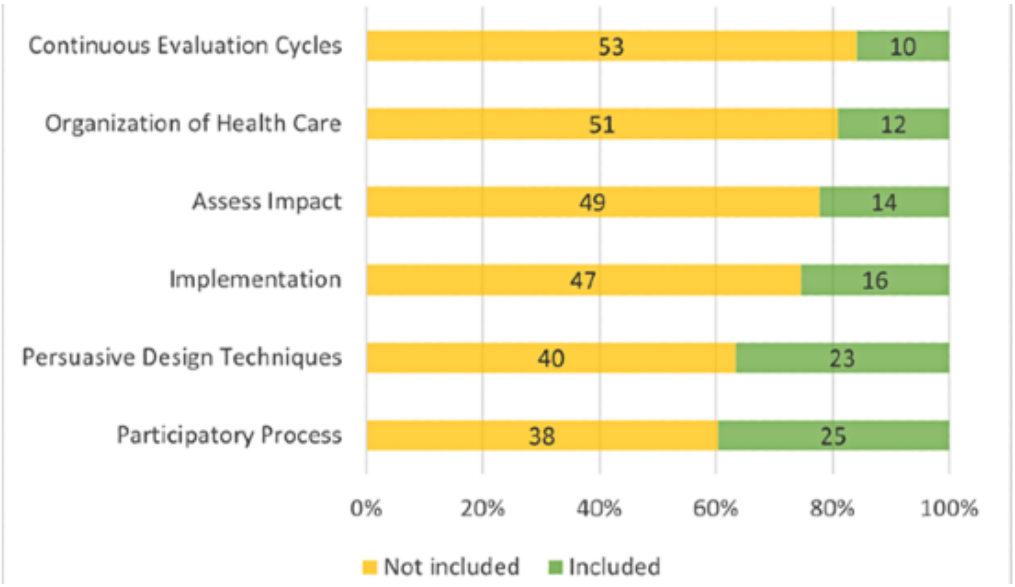


Main Findings

To What Extent Does the Available Literature Discuss Features That Promote the Use of Contact Tracing Apps by Interested Parties?

Overall, all of the six principles of the Framework to Improve the Uptake and Impact of eHealth Technologies [41] were covered in less than half of the selected articles. Among these, mentions of either use cases or recommendations for a participatory process in the development of contact tracing apps appeared in 40% (n=25) of the 63 articles, which made it the most frequently discussed component (Figure 4). Conversely, only 16% (n=10) of the studies discussed continuous evaluation cycles of contact tracing apps. Persuasive design techniques were mentioned in 37% (n=23) of the articles, followed by implementation considerations (n=16, 25%), advanced methods to assess impacts (n=14, 22%), and foresight of changes in the organization of health care (n=12, 19%).

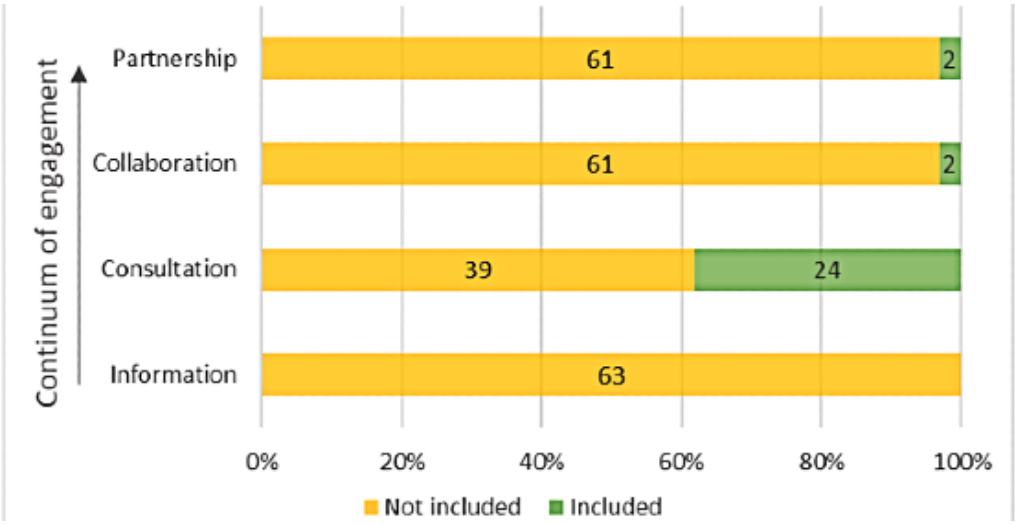
Figure 4. Proportion of articles integrating holistic eHealth development principles.



How Have Patients and Citizens Been Engaged in the Design and Implementation of These Apps?

None of the selected articles referred to documentation given to citizens or patients on COVID-19 contact tracing apps, leaving the Montreal model "information" component [42] unaddressed (Figure 5). On the other hand, stakeholder consultation was the most frequently described level of engagement (n=24, 38%). It is worth noting that only 1 study related to patients and only 1 mentioned the creation of a focus group (in contrast to surveys of the public and reviews of public commentaries on the internet). One article briefly mentioned collaboration, as the responses to the conducted surveys "likely prompted the pivot to wearable tech due to its observation of the public's reluctance to use their mobile phones for contact-tracing" [50]. Only 3% (2/63) mentioned the importance of partnership in the development and implementation of contact tracing apps, but none of the included articles described an app development process occurring in partnership with citizens or patients.

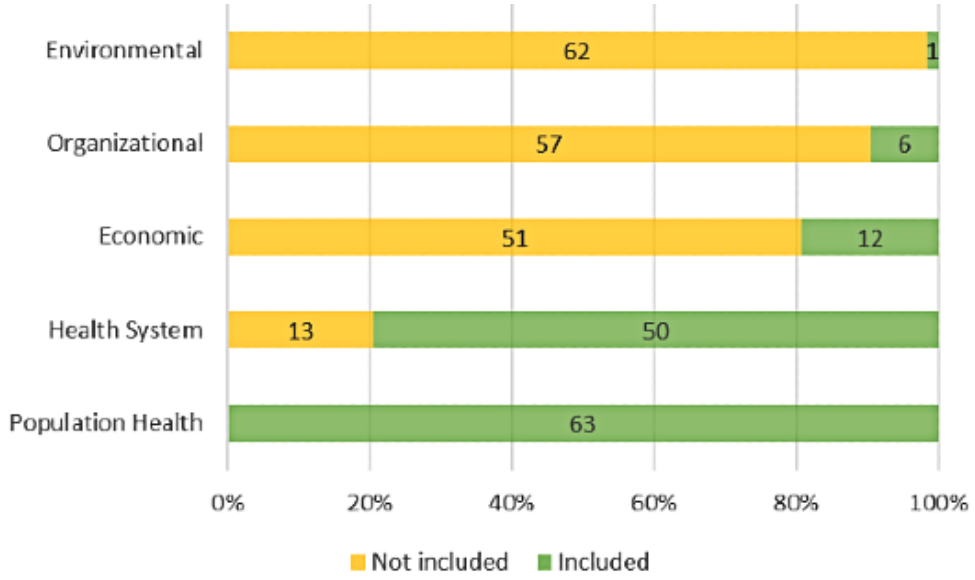
Figure 5. Proportion of articles integrating each level along the patient or citizen engagement continuum.



Does the Development of These Apps Correspond to Responsible Research and Innovation?

Among the five domains of the Responsible Innovation in Health Assessment Tool [46], population health appeared in all of the selected studies, constituting the most frequently addressed component (Figure 6). In contrast, only 2% (n=1) of the 63 studies discussed the environmental component. A discussion of the health system was included in 79% (n=50) of the articles, while the economic and organizational dimensions were present in 19% (n=12) and 10% (n=6), respectively.

Figure 6. Proportion of articles integrating Responsible Innovation in Health domains.



Design Considerations and Implementation Recommendations

The principal considerations and recommendations regarding contact tracing apps in the COVID-19 pandemic context were drawn from the key findings of analyzed articles and are summarized in Table 1. Our results show that the main themes of the 162 theoretical considerations and practical recommendations include data management (n=58, 35.8%), user experience (n=48, 29.6%), communication (n=23, 14.2%), research and implementation methodology (n=18, 11.1%), and the engagement of stakeholders (n=15, 9.3%). Here, user experience refers to the main reactions or perceived barriers that may affect uptake (which mainly comes from survey results), incentivization, app functionality, and other design considerations. Data management designates app architecture, data protection, privacy, monitoring of data, laws governing the use of data, and ethical consideration of data use. Communication covers all recommendations that communication from governments, health authorities, or app developers must be clear and transparent or other measures that may increase public trust in those institutions, such as rebranding and optimizing already existing apps, and reframing the terminology used. The engagement of stakeholders refers to consultation with citizens or patients, strategies to reach vulnerable populations, and theoretical considerations about equity in health access. Finally, research and implementation methodology includes the frameworks proposed to guide future research, general considerations of how to improve uptake, and theoretical considerations of what constitutes ethical app use.

Table 1. Themes of considerations and recommendations regarding app design and implementation.

Themes	Considerations and recommendations found in the literature (N=162), n (%)
Data management (eg, the use of blockchain [53-56] and sunset clause [57])	58 (35.8)
User experience (eg, rewards to app users [58] and simple design and user interface for interaction [59])	48 (29.6)
Communication (eg, clarifying false beliefs about the app [29] and prevention campaign about individual risks [60])	23 (14.2)
Research and implementation methodology (eg, NPT[a] framework to guide development and evaluation of complex DPT[b] interventions [61], apps must be necessary, proportional, scientifically valid, and time bound [62])	18 (11.1)
Engagement of stakeholders (eg, low-cost wristband in low-socioeconomic areas [54] or alignment of the app with local culture and vulnerable populations [63])	15 (9.3)

[a]NPT: Normalization Process Theory.

[b]DPT: digital proximity tracing.

Discussion

Principal Findings

This study demonstrates that the volume of articles on design and implementation considerations for COVID-19 contact tracing apps grew rapidly in the first months of the pandemic and peaked in August 2020. As the evidence base increased, the proportion of opinion pieces decreased in comparison to surveys and review articles. This is not surprising given the novelty of this technology and the urgency around technological developments triggered by the pandemic. Accordingly, earlier articles mostly discussed principles related to the development and implementation of contact tracing apps mostly in theoretical terms, while the proportion of empirical articles increased

starting in July. Despite the growing number of publications, their scope remained limited with respect to design and implementation considerations. Our findings demonstrate that critical design and implementation considerations were lacking in the early academic literature on contact tracing apps.

In fact, less than half of the selected articles cited the need for a participatory process, which nonetheless constituted the most frequently referenced item of the Framework to Improve the Uptake and Impact of eHealth Technologies. Only 40% (25/63) of the articles presented evidence of public or patient engagement in the contact tracing app development process. This stands in stark contrast with the ideal of "co-creation from ideation to operationalization" in eHealth technologies, as described in this framework [41]. Other principles from the same framework were infrequently addressed, such as the need for continuous evaluation cycles, the creation of new processes for health care delivery, the need to assess impact, and the proposing of specific actions for implementation.

We further analyzed the level of citizen and patient engagement in the development of contact tracing apps by using the Montreal model [42]. We found that, among the 28 studies that addressed this component, nearly all (n=24, 86%) consisted of public consultations. Although essential to inform design and implementation of eHealth technologies, consultation ranks lower on the patient engagement continuum as compared to collaboration and partnership. This imbalance may have resulted from the urgency of the public health crisis combined with the perception that citizen and patient engagement is time-consuming [64-66]. However, the upstream efforts deployed to promote collaboration and partnership with users provided key insights at early stages of development [67,68]. Ignoring user engagement or relying solely on downstream consultation may in fact delay implementation by triggering a need to redesign certain features or even contribute to ultimately unsuccessful ventures [68,69].

Moreover, when we assessed these studies according to the Responsible Innovation in Health framework [46], we found that most of the 63 articles described the impact of contact tracing apps at the population level and on health systems (n=63, 100% and n=50, 79%, respectively), mainly in theoretical terms. However, only a few acknowledged the environmental, organizational, and economic value domains (each domain was found in less than 20% of the included studies). Their near absence from the literature may indicate that these domains are not perceived as being relevant to the development of COVID-19 contact tracing apps.

The incompleteness of the academic literature on the design and implementation characteristics of COVID-19 contact tracing apps stems in part from the novelty of the topic and the need for timely innovations to fight the pandemic. As such, we did not expect the literature to address all the domains of the selected frameworks. However, few academic publications highlighted the need to assess the impact of this intervention or consider organizational challenges related to its deployment, among other key elements to successful and responsible eHealth innovation. Likewise, the limited references to higher forms of upstream end user engagement stand in contrast to this technology's inherent reliance on widespread adoption. It may be that app developers or researchers did not perceive some of these factors as bearing sufficient relevance for incorporation at the development stage or for later publication. Nonetheless, more than one year since the start of the pandemic, many of these elements are still lacking and were highlighted as essential to the success of COVID-19 contact tracing apps [38]. Considering these findings, we hypothesize that the sense of urgency instilled by the pandemic motivated shortcuts away from a full compliance with best practices in eHealth development. In turn, this may have contributed to the unconvincing implementation of contact tracing apps to date. Future research will help determine which factors are most associated with the development of sustainable, feasible, ethically acceptable, and socially desirable technological solutions in the context of a public health emergency. As specific evidence on this topic continues to accumulate, we draw upon the existing literature to make some practical recommendations for successful contact tracing apps and related eHealth innovations.

Recommendations Regarding the Design Considerations and Implementation of COVID-19 Contact Tracing Apps

Data Management

Data protection and the looming risk of mass surveillance have understandably dominated the debate on the ethical, legal, and social implications of contact tracing apps. Users must therefore be intelligibly informed of the steps taken to protect their privacy. One consideration is that the risk of the proposed contact tracing app should be compared with that of the frequently used apps that most people have on their cellphones. For example, the New York Times "Privacy Project" recently revealed the tracking of millions of unsuspecting Americans through location-sharing apps (eg, apps used to access directions, weather information, or local fidelity programs) [70]. The heightened public scrutiny of contact tracing apps may thus serve as an opportunity to promote digital literacy and reflect on the responsible use (or misuse) of digital tools.

Communication

Clear and transparent messaging developed with and for citizens can be promoted through simple and familiar means, such as concise information labels inspired by those of the food industry. This approach was used by members of the health care machine learning community to promote the transparent and responsible use of clinical decision support tools [71]. If they are shown to be effective, clear communication and efforts to improve health literacy will further promote stakeholders' engagement [72,73] and may increase public trust in governments and health authorities. Rebranding or reframing such as changing the terminology from tracking apps to exposure notification apps may also be a solution that would improve both public trust and app uptake. Moreover, given the large number of contact tracing apps developed thus far, a consistent presentation of fundamental design components (see the Introduction section) would assist stakeholders in their evaluation of a given app. It may also promote efficient comparisons between different apps and perhaps limit the risk of unnecessary duplication.

Engaging Key Stakeholders

Key stakeholders of contact tracing apps include potential users, technology developers, policy makers, and funding agencies. As previously noted, the prevailing form of engagement has been through public consultations, which is recommended in the analyzed articles. We would therefore argue that app developers and the various institutions that implement such technologies would benefit from greater upstream collaboration and partnership with individuals from diverse backgrounds, including patients. Engaging key stakeholders early in the process will also help identify the right problem and constraints, eventually narrowing the range of most suitable technological tools. Indeed, even if a contact tracing app perfectly identified high-risk contacts, it would not achieve the desired outcome of reduced viral transmission if it required prohibitively expensive hardware or if it relied on massive viral testing in a strained health care system unable to provide a sufficient number of diagnostic tests.

As previously described, the potential risks of contact tracing apps may disproportionately affect minority and marginalized groups [35-37]. It is important to seek such participation in the development of these tools to mitigate and, ideally, eliminate the risk of increasing disparities through their use. This recommendation is also supported by the literature, which prescribes strategies for reaching low socioeconomic groups and older adults. Unfortunately, it may be difficult to engage with certain groups, such as for people who experience homelessness, (digital) illiteracy, living in remote settings, or underrepresentation within the typical channels of citizen engagement due to systemic racism or other structural vulnerabilities. In these cases, one practical option by which app developers could enhance the diversity of voices contributing to their product is to partner with community leaders and individuals who work closely with underrepresented groups. Although this will require an early investment of time and resources, the dividends in terms of greater inclusiveness, accurate problem identification, and a broader assessment of the impact on outcomes across different populations will strengthen a project's chances of achieving responsible design and successful implementation. Public and private funders can also play an important role in ensuring that effective collaboration occurs early in the development of contact tracing apps by prioritizing proposals that promote a participatory process.

User Experience

A simple and intuitive user interface may not only enhance the user experience at an individual level, it may also improve uptake at the population level. Apps that focus on interactive design features can help users better understand how to use the app correctly and more effectively, whereas apps that are focused on information may reduce the assimilation of this information by users if they are not visually appealing [74]. Moreover, resolving in-app technical shortcomings will likely improve the user experience. In fact, an Australian survey noted that 24% of respondents had listed technical concerns as a reason not to download a contact tracing app during the COVID-19 pandemic [27]. An easy system for reporting technical difficulties and a minimally disruptive evaluation of other app functionalities may therefore increase use. Moreover, since one main barrier to app uptake is privacy and data security concerns [21-23,27,29,49,75,76], users can be empowered by designs that use customizable app functionalities. For example, app users could be offered choices in the technology used to collect the data (GPS or Bluetooth), how the data is stored [77], and the level of interaction with local health agencies.

Simple and useful apps will likely incentivize uptake and use. This can be supported by developing apps that require readily available personal technology, personalized updates on pertinent and accurate information aligned with local guidelines, minimal disruption to daily functioning (eg, minimizing battery use [78]), and facilitated testing in collaboration with local health care systems. Poor coordination with local health actors may therefore significantly compromise the potential user value of contact tracing apps. On the other hand, locally integrated apps may help streamline testing when it is indicated and provide current trustworthy recommendations. In many instances, COVID-19 contact tracing apps were designed to complement established traditional public health interventions. Apps that work in silos may therefore be less appealing to the public, as their perceived usefulness and effectiveness may be compromised [79]. Furthermore, implementation should aim for interoperability in the identification and notification of high-risk contacts in the greater interest of public health. This may be more attainable in jurisdictions that share compatible legal frameworks for data protection and privacy, such as the General Data Protection Regulation [80].

Financial incentives should also be considered. A study that focused on tracking the use of Germany's official contact tracing app, Corona-Warn-App, found that app uptake is more prevalent among older populations, individuals with pre-existing conditions, and those with high levels of education and income. Additionally, the study reported that information interventions, in the form of short videos addressing privacy, effectiveness, and app functionality issues, were useful in increasing users' knowledge about the app but were not effective in driving uptake. On the other hand, interventions that provide a monetary incentive (as low as €1 [US $1.22], €3 [US $3.67], or €5 [US $6.11]) upon installation were found to be useful in increasing uptake [81].

Research and Implementation Strategies

Empirical evidence is urgently needed to determine whether the benefits of a given contact tracing app significantly outweigh its risks. Although contact tracing technologies were used in prior Ebola and influenza outbreaks, there is limited empirical evidence on their effectiveness. As for their impact in reducing viral transmission during the COVID-19 pandemic, it is mainly based on mathematical simulation models that used varied assumptions and methodologies [18]. Real-life estimated treatment effects, beyond the current simulation models, are urgently needed, with precise descriptions of the context in which the tool was used. In addition to the estimated effect of contact tracing apps on public health outcomes (eg, the impact on the basic reproduction number R_0), studies should also report on the input data and the corresponding outputs. At the least,

information on app downloading and daily use must be made available, along with a description of key app features. Such information would enable the study of factors that enhance uptake and the relationship between app uptake and public health effectiveness. Observational studies on contact tracing app effectiveness during a period when large-scale interventions are continuously being proposed and deimplemented will certainly be limited by significant biases. They can nonetheless shed light on the role played by digital contact tracing during this pandemic and future infectious disease outbreaks. This observational evidence can then be compared to outcomes from simulation-based studies, including a recent modeling that suggests that a digital contact tracing and exposure notification system can support traditional public health interventions in reducing transmission, even at participation levels as low as 15% [82].

In addition, academic reports on contact tracing apps and related eHealth innovations must adhere to established reporting guidelines such as CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) for eHealth and mHealth interventions [83]. The global scale of contact tracing apps makes the incorporation of the key domains and subdomains proposed in the Responsible Innovation in Health framework even more crucial than in local innovations. eHealth proposals should therefore indicate their predicted or estimated impact on health inequalities, health system inclusiveness and responsiveness, frugality, business model, and eco-responsibility. Funding agencies and academic journals must request that submissions comprehensively address these issues.

Limitations

The findings of this scoping review must be interpreted considering certain limitations. First, this study includes articles published up to November 6, 2020. As the volume of publications continues to grow, a systematic review focused on narrower questions related to contact tracing apps may become relevant. We attempted to maximize the reach of our search by including multiple databases and by developing a rigorous study selection process. Relevant articles only available in the gray literature or exclusively in governmental databases may nevertheless have been missed, although official apps developed by governments, such as the National Health Service contact tracing app, were included in several of the articles analyzed. Furthermore, given the purposefully broad question posed by this study, the reproducibility of data extraction and charting presented some challenges. We addressed them by relying on well-established and complementary frameworks that were particularly appropriate to the research question. Moreover, multiple authors reviewed the process and provided supportive statements when a particular component or domain was considered to be present in a given article.

Conclusions

The emerging academic literature on contact tracing apps reveals significant knowledge gaps regarding their design and implementation. Key stakeholders are thus limited in their ability to critically appraise this eHealth innovation. Most of the included studies lacked fundamental aspects of the successful eHealth development and implementation framework. Similarly, few articles described the impact of contact tracing apps on the environmental, organizational, and economic domains, which are essential to evaluate responsible innovation in health. Among the studies that described a form of public participation, nearly all of them relied on consultation as opposed to collaboration or partnership. These overlooked components of eHealth development and implementation may have contributed to the modest uptake of contact tracing tools worldwide. They suggest a critical gap between theory and practice, whereby numerous academic sources promote a holistic and participatory approach to eHealth innovation, but few products incorporate them. Partnerships between app developers, researchers, policy makers, and users early in the development process will narrow this gap. Transparent, systematic, and comprehensive reporting of COVID-19 contact tracing app outcomes will further enable their critical appraisal. The lessons learned about the social acceptability of contact tracing apps as they were deployed at an unprecedented pace and scale must serve in future iterations of this innovation and in the development of other eHealth technologies aimed at sustainably supporting public health. They must attest to the importance of stakeholder engagement, problem identification, minimal system disruptions, longitudinal outcome measurement, and use incentivization.

Acknowledgments

This research was supported by the Observatoire international sur les impacts sociétaux de l'IA et du numérique and the Fonds de recherche du Québec. MPP also receives funding from the Fonds de recherche du Québec pour le soutien à la bourse de chercheur senior. We would like to thank Eleonora Bogdanova for her invaluable help in the literature search and Olivier Demers-Payette, Nathalie de Marcellis-Warin, and Christophe Mondin for their technical and methodological insight.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) Checklist.
[PDF File (Adobe PDF File), 110 KB - mhealth_v9i6e27102_app1.pdf]

Multimedia Appendix 2
Additional database search strategies.
[PDF File (Adobe PDF File), 64 KB - mhealth_v9i6e27102_app2.pdf]

Multimedia Appendix 3
Database of included articles and analysis.
[XLSX File (Microsoft Excel File), 41 KB - mhealth_v9i6e27102_app3.xlsx]

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
2. COVAX: working for global equitable access to COVID-19 vaccines. World Health Organization. URL: https://www.who.int/initiatives/act-accelerator/covax [accessed 2021-01-04]
3. New COVID-19 variants. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/transmission/variant.html [accessed 2021-01-04]
4. Coronavirus disease 2019 (COVID-19) : situation report, 73. World Health Organization. 2020 Apr 02. URL: https://apps.who.int/iris/handle/10665/331686 [accessed 2020-12-02]
5. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic transmission, the Achilles' heel of current strategies to control Covid-19. N Engl J Med 2020 May 28;382(22):2158-2160 [FREE Full text] [doi: 10.1056/NEJMe2009758] [Medline: 32329972]
6. Tokar S. UCSF Team presents ethical framework for effective COVID-19 contact tracing. UC San Francisco. URL: https://www.ucsf.edu/news/2020/10/418841/ucsf-team-presents-ethical-framework-effective-covid-19-contact-tracing [accessed 2020-12-02]
7. The impact of chronic underfunding on America’s public health system: trends, risks, and recommendations, 2020. Trust for America's Health. 2020. URL: https://www.tfah.org/report-details/publichealthfunding2020/ [accessed 2020-12-02]
8. Echelard J, Méthot F, Nguyen H, Pomey M. Medical student training in eHealth: scoping review. JMIR Med Educ 2020 Sep 11;6(2):e20027 [FREE Full text] [doi: 10.2196/20027] [Medline: 32915154]
9. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. Lancet 2020 Aug;2(8):e435-e440.
10. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
11. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-Peer Contact Tracing: Development of a Privacy-Preserving Smartphone App. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
12. Parker MJ, Fraser C, Abeler-Dörner L, Bonsall D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. J Med Ethics 2020 Jul;46(7):427-431 [FREE Full text] [doi: 10.1136/medethics-2020-106314] [Medline: 32366705]
13. Mondin C, de Marcellis-Warin N. Recension des solutions technologiques développées dans le monde afin de limiter la propagation de la COVID-19 et typologie des applications de traçage. DocDroid. 2020 Oct. URL: https://tinyurl.com/vrxkmhym [accessed 2020-12-02]
14. Exposure notifications: using technology to help public health authorities fight COVID-19. Google. URL: https://www.google.com/intl/en_us/covid19/exposurenotifications/ [accessed 2020-12-02]
15. Download COVID Alert today. Government of Canada. 2020 Sep 15. URL: https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert.html [accessed 2020-12-09]
16. Exit through the App Store? A rapid evidence review on the technical considerations and societal implications of using technology to transition from the COVID-19 crisis. Ada Lovelace Institute. 2020 Apr 20. URL: https://tinyurl.com/kp3v38c7 [accessed 2020-12-09]
17. Kahn J, editor. Digital Contact Tracing for Pandemic Response: Ethics and Governance Guidance. Baltimore, MD: Johns Hopkins University Press; 2020.
18. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. Lancet 2020 Nov;2(11):e607-e621.
19. Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown. University of Oxford: Our Research. URL: https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown [accessed 2021-04-16]
20. O'Neill PH. No, coronavirus apps don’t need 60% adoption to be effective. MIT Technology Review. URL: https://www.technologyreview.com/2020/06/05/1002775/covid-apps-effective-at-less-than-60-percent-download/ [accessed 2021-04-16]

21. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
22. Bachtiger P, Adamson A, Quint JK, Peters NS. Belief of previous COVID-19 infection and unclear government policy are associated with reduced willingness to participate in app-based contact tracing: a UK-wide observational study of 13,000 patients. medRxiv. Preprint posted online on June 6, 2020. [doi: 10.1101/2020.06.03.20120337]
23. Jansen-Kosterink S, Hurmuz M, den OM, van VL. Predictors to use mobile apps for monitoring COVID-19 symptoms and contact tracing: a survey among Dutch citizens. medRxiv. Preprint posted online on June 2, 2020. [doi: 10.1101/2020.06.02.20113423]
24. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mölken M. COVID-19 contact tracing apps: predicted uptake in the Netherlands based on a discrete choice experiment. JMIR Mhealth Uhealth 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: 10.2196/20741] [Medline: 32795998]
25. Kaspar K. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: quantitative survey study. J Med Internet Res 2020 Aug 27;22(8):e21613 [FREE Full text] [doi: 10.2196/21613] [Medline: 32759100]
26. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. Ir J Med Sci 2020 Oct 16:1-25 [FREE Full text] [doi: 10.1007/s11845-020-02389-y] [Medline: 33063226]
27. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. More than privacy: Australians' concerns and misconceptions about the COVIDSafe App. medRxiv. Preprint posted online on July 29, 2020. [doi: 10.1101/2020.06.09.20126110]
28. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a Health Belief Model approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
29. Montagni I, Roussel N, Thiébaut R, Tzourio C. The French Covid-19 contact tracing app: usage and opinions by students in the health domain. medRxiv. Preprint posted online on November 12, 2020. [doi: 10.1101/2020.10.23.20218214]
30. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Are COVID-19 proximity tracing apps working under real-world conditions? Indicator development and assessment of drivers for app (non-) use. medRxiv. Preprint posted online on September 2, 2020. [doi: 10.1101/2020.08.29.20184382]
31. O'Neill PH, Ryan-Mosley T, Johnson B. A flood of coronavirus apps are tracking us: now it's time to keep track of them. MIT Technology Review. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/ [accessed 2021-04-16]
32. MIT Technology Review. Covid Tracing Tracker. Google Docs. URL: https://docs.google.com/spreadsheets/d/1ATalASO8KtZMx__zJREoOvFh0nmB-sAqJ1-CjVRSCOw/edit#gid=1464910624 [accessed 2021-04-08]
33. Flood CM, MacDonnell V, Thomas B, Wilson K. Reconciling civil liberties and public health in the response to COVID-19: an RSC policy briefing. The Royal Society of Canada. 2020 Sep. URL: https://rsc-src.ca/sites/default/files/CL%20PB_EN.pdf [accessed 2020-12-02]
34. Veinot T, Mitchell H, Ancker J. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 01;25(8):1080-1088 [FREE Full text] [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
35. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet 2020 Aug;2(8):e425-e434.
36. Hendl T, Chung R, Wild V. Pandemic surveillance and racialized subpopulations: mitigating vulnerabilities in COVID-19 apps. J Bioeth Inq 2020 Dec;17(4):829-834 [FREE Full text] [doi: 10.1007/s11673-020-10034-7] [Medline: 32840858]
37. Toh A, Brown D. How digital contact tracing for covid-19 could worsen inequality. Human Rights Watch. 2020 Jun 04. URL: https://www.hrw.org/news/2020/06/04/how-digital-contact-tracing-covid-19-could-worsen-inequality [accessed 2020-12-02]
38. Colizza V, Grill E, Mikolajczyk R, Cattuto C, Kucharski A, Riley S, et al. Time to evaluate COVID-19 contact-tracing apps. Nat Med 2021 Mar;27(3):361-362. [doi: 10.1038/s41591-021-01236-6] [Medline: 33589822]
39. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
40. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
41. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
42. Pomey M, Flora L, Karazivan P, Dumez V, Lebel P, Vanier MC, et al. [The Montreal model: the challenges of a partnership relationship between patients and healthcare professionals]. Sante Publique 2015;27(1 Suppl):S41-S50. [Medline: 26168616]

43. Guide d'implantation du partenariat de soins et de services: vers une pratique collaborative optimale entre intervenants et avec le patient. Formation RUISSS. 2014 Feb 10. URL: http://ena.ruis.umontreal.ca/pluginfile.php/256/coursecat/description/Guide_implantation1.1.pdf [accessed 2020-12-02]
44. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. Health Aff (Millwood) 2013 Feb;32(2):223-231. [doi: 10.1377/hlthaff.2012.1133] [Medline: 23381514]
45. Vanier MC, Dumez V, Drouin E, Brault I, MacDonald SA, Boucher A, et al. Partners in interprofessional education: integrating patients-as-trainers. 2014 Presented at: Partnering with Patients, Families, and Communities to Link Interprofessional Practice and Education; April 2014; Arlington, VA URL: http://macyfoundation.org/docs/macypubs/JMF_PartneringwithPFC.pdf
46. Pacifico Silva H, Lehoux P, Miller FA, Denis J. Introducing responsible innovation in health: a policy-oriented framework. Health Res Policy Syst 2018 Sep 10;16(1):90 [FREE Full text] [doi: 10.1186/s12961-018-0362-5] [Medline: 30200985]
47. Loi M. How to fairly incentivise digital contact tracing. J Med Ethics 2020 Jul 09:medethics-2020-106388 [FREE Full text] [doi: 10.1136/medethics-2020-106388] [Medline: 32647047]
48. Nature. Show evidence that apps for COVID-19 contact-tracing are secure and effective. Nature 2020 Apr;580(7805):563. [doi: 10.1038/d41586-020-01264-1] [Medline: 32350479]
49. Williams SN, Armitage C, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. PsyArXiv. Preprint posted online on May 15, 2020. [doi: 10.31234/osf.io/ra93b]
50. Lee T, Lee H. Tracing surveillance and auto-regulation in Singapore: 'smart' responses to COVID-19. Media Int Aust 2020 Aug 12;177(1):47-60. [doi: 10.1177/1329878x20949545]
51. 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] [Medline: 32609622]
52. Nijsingh N, van Bergen A, Wild V. Applying a precautionary approach to mobile contact tracing for COVID-19: the value of reversibility. J Bioeth Inq 2020 Dec;17(4):823-827 [FREE Full text] [doi: 10.1007/s11673-020-10004-z] [Medline: 32840831]
53. Arifeen MM, Al Mamun A, Kaiser MS, Mahmud M. Blockchain-enable contact tracing for preserving user privacy during COVID-19 outbreak. Preprints. Preprint posted online on July 22, 2020. [doi: 10.20944/preprints202007.0502.v1]
54. Chowdhury MJM, Ferdous MS, Biswas K, Chowdhury N, Muthukkumarasamy V. COVID-19 contact tracing: challenges and future directions. Preprints. Preprint posted online on June 19, 2020. [doi: 10.20944/preprints202006.0240.v2]
55. Klaine PV, Zhang L, Zhou B, Sun Y, Xu H, Imran M. Privacy-preserving contact tracing and public risk assessment using blockchain for COVID-19 pandemic. IEEE Internet Things Magazine 2020 Sep;3(3):58-63. [doi: 10.1109/iotm.0001.2000078]
56. Xu H, Zhang L, Onireti O, Fang Y, Buchanan WJ, Imran MA. BeepTrace: blockchain-enabled privacy-preserving contact tracing for COVID-19 pandemic and beyond. IEEE Internet Things J 2021 Mar 1;8(5):3915-3929. [doi: 10.1109/jiot.2020.3025953]
57. Culnane C, Leins K, Rubinstein BI. Tracking, tracing, trust: contemplating mitigating the impact of COVID-19 through technological interventions. Med J Aust 2020 Jul;213(2):94-94.e1 [FREE Full text] [doi: 10.5694/mja2.50683] [Medline: 32583470]
58. Cruz MM, Oliveira RS, Beltrão APV, Lopes PHB, Viterbo J, Trevisan DG, et al. Assessing the level of acceptance of a crowdsourcing solution to monitor infectious diseases propagation. 2020 Presented at: 2020 IEEE International Smart Cities Conference; September 28-October 1, 2020; Piscataway, NJ. [doi: 10.1109/isc251055.2020.9239069]
59. Azad MA, Arshad J, Akmal SMA, Riaz F, Abdullah S, Imran M, et al. A first look at privacy analysis of COVID-19 contact tracing mobile applications. IEEE Internet Things J 2020:1-1. [doi: 10.1109/jiot.2020.3024180]
60. Guillon M, Kergall P. Attitudes and opinions on quarantine and support for a contact-tracing application in France during the COVID-19 outbreak. Public Health 2020 Nov;188:21-31 [FREE Full text] [doi: 10.1016/j.puhe.2020.08.026] [Medline: 33059232]
61. von Wyl V. Challenges for nontechnical implementation of digital proximity tracing during the COVID-19 pandemic: media analysis of the SwissCovid app. JMIR Mhealth Uhealth 2021 Feb 26;9(2):e25345 [FREE Full text] [doi: 10.2196/25345] [Medline: 33606658]
62. Morley J, Cowls J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. Nature 2020 Jun;582(7810):29-31. [doi: 10.1038/d41586-020-01578-0] [Medline: 32467596]
63. Vandamme A, Nguyen T. Belgium - concerns about coronavirus contact-tracing apps. Nature 2020 May;581(7809):384. [doi: 10.1038/d41586-020-01552-w] [Medline: 32457527]
64. Gross O. Faire de l'engagement des patients une réalité. In: L'engagement des patients au service du système de santé. London: Doin - John Libbey Eurotext; 2018.
65. Martineau JT, Minyaoui A, Boivin A. Partnering with patients in healthcare research: a scoping review of ethical issues, challenges, and recommendations for practice. BMC Med Ethics 2020 May 11;21(1):34 [FREE Full text] [doi: 10.1186/s12910-020-0460-0] [Medline: 32393230]

66. Ellis LE, Kass NE. How are PCORI-funded researchers engaging patients in research and what are the ethical implications? AJOB Empir Bioeth 2017;8(1):1-10. [doi: 10.1080/23294515.2016.1206045] [Medline: 28949867]
67. Subramanyam R, Weisstein FL, Krishnan MS. User participation in software development projects. Commun ACM 2010 Mar 01;53(3):137-141. [doi: 10.1145/1666420.1666455]
68. Supple D, Roberts A, Hudson V, Masefield S, Fitch N, Rahmen M, U-BIOPRED PIP group. From tokenism to meaningful engagement: best practices in patient involvement in an EU project. Res Involv Engagem 2015;1:5 [FREE Full text] [doi: 10.1186/s40900-015-0004-9] [Medline: 29062494]
69. Boote J, Dalgleish M, Freeman J, Jones Z, Miles M, Rodgers H. 'But is it a question worth asking?' A reflective case study describing how public involvement can lead to researchers' ideas being abandoned. Health Expect 2014 Jun;17(3):440-451 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00771.x] [Medline: 22646745]
70. Thompson SA, Warzel C. Twelve million phones, one dataset, zero privacy. The New York Times. 2019 Dec 19. URL: https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html [accessed 2021-01-11]
71. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. NPJ Digit Med 2020;3:41. [doi: 10.1038/s41746-020-0253-3] [Medline: 32219182]
72. Maurer M, Dardess P, Carman KL, Frazier K, Smeeding L. Guide to patient and family engagement: environmental scan report. Agency for Healthcare Research and Quality. 2012 May. URL: https://www.ahrq.gov/sites/default/files/publications/files/ptfamilyscan.pdf [accessed 2021-01-07]
73. Kaisler RE, Missbach B. Co-creating a patient and public involvement and engagement 'how to' guide for researchers. Res Involv Engagem 2020;6:32 [FREE Full text] [doi: 10.1186/s40900-020-00208-3] [Medline: 32566249]
74. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A systematic review of smartphone applications available for corona virus disease 2019 (COVID19) and the assessment of their quality using the Mobile Application Rating Scale (MARS). J Med Syst 2020 Aug 10;44(9):164 [FREE Full text] [doi: 10.1007/s10916-020-01633-3] [Medline: 32779002]
75. Sharma S, Singh G, Sharma R, Jones P, Kraus S, Dwivedi YK. Digital health innovation: exploring adoption of COVID-19 digital contact tracing apps. IEEE Trans Eng Manage 2020:1-17. [doi: 10.1109/tem.2020.3019033]
76. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]
77. Klar R, Lanzerath D. The ethics of COVID-19 tracking apps – challenges and voluntariness. Res Ethics 2020 Aug 05;16(3-4):1-9. [doi: 10.1177/1747016120943622]
78. Active Irish COVID-19 tracing app users drop on battery problem - HSE. Reuters. 2020 Aug 27. URL: https://www.reuters.com/article/us-health-coronavirus-ireland-apps-idUSKBN25N1PA [accessed 2021-01-11]
79. Carman KL, Workman TA. Engaging patients and consumers in research evidence: applying the conceptual model of patient and family engagement. Patient Educ Couns 2017 Jan;100(1):25-29 [FREE Full text] [doi: 10.1016/j.pec.2016.07.009] [Medline: 27473639]
80. GDPR checklist for data controllers. General Data Protection Regulation. URL: https://gdpr.eu/checklist/ [accessed 2020-04-13]
81. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
82. Abueg M, Hinch R, Wu N, Liu L, Probert W, Wu A, et al. Modeling the combined effect of digital exposure notification and non-pharmaceutical interventions on the COVID-19 epidemic in Washington state. medRxiv. Preprint posted online on September 2, 2020. [doi: 10.1101/2020.08.29.20184135]
83. 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] [Medline: 22209829]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

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

Edited by L Buis; submitted 14.01.21; peer-reviewed by W Wilkowska, D Leightley, K Eason; comments to author 24.03.21; revised version received 14.04.21; accepted 20.05.21; published 09.06.21.

Please cite as:
Osmanlliu E, Rafie E, Bédard S, Paquette J, Gore G, Pomey MP
Considerations for the Design and Implementation of COVID-19 Contact Tracing Apps: Scoping Review
JMIR Mhealth Uhealth 2021;9(6):e27102
URL: https://mhealth.jmir.org/2021/6/e27102
doi:10.2196/27102
PMID:34038376

Review

State of the Art in Adoption of Contact Tracing Apps and Recommendations Regarding Privacy Protection and Public Health: Systematic Review

Katarzyna Kolasa[1], PhD; Francesca Mazzi[2,3], LLM; Ewa Leszczuk-Czubkowska[4], MSc; Zsombor Zrubka[5,6], MD, MBA, PhD; Márta Péntek[5], Prof Dr, PhD

[1]Division of Health Economics and Healthcare Management, Kozminski University, Warsaw, Poland

[2]Queen Mary University of London, London, United Kingdom

[3]Maastricht University, Maastricht, Netherlands

[4]Kozminski University, Warsaw, Poland

[5]Health Economics Research Center, Óbuda University, Budapest, Hungary

[6]Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Budapest, Hungary

Corresponding Author:
Katarzyna Kolasa, PhD
Division of Health Economics and Healthcare Management
Kozminski University
57/59 Jagiellonska St
Warsaw, 03-301
Poland
Phone: 48 22 519 21 00
Email: kkolasa@kozminski.edu.pl

Abstract

Background: During the COVID-19 pandemic, contact tracing apps have received a lot of public attention. The ongoing debate highlights the challenges of the adoption of data-driven innovation. We reflect on how to ensure an appropriate level of protection of individual data and how to maximize public health benefits that can be derived from the collected data.

Objective: The aim of the study was to analyze available COVID-19 contact tracing apps and verify to what extent public health interests and data privacy standards can be fulfilled simultaneously in the process of the adoption of digital health technologies.

Methods: A systematic review of PubMed and MEDLINE databases, as well as grey literature, was performed to identify available contact tracing apps. Two checklists were developed to evaluate (1) the apps' compliance with data privacy standards and (2) their fulfillment of public health interests. Based on both checklists, a scorecard with a selected set of minimum requirements was created with the goal of estimating whether the balance between the objective of data privacy and public health interests can be achieved in order to ensure the broad adoption of digital technologies.

Results: Overall, 21 contact tracing apps were reviewed. In total, 11 criteria were defined to assess the usefulness of each digital technology for public health interests. The most frequently installed features related to contact alerting and governmental accountability. The least frequently installed feature was the availability of a system of medical or organizational support. Only 1 app out of 21 (5%) provided a threshold for the population coverage needed for the digital solution to be effective. In total, 12 criteria were used to assess the compliance of contact tracing apps with data privacy regulations. Explicit user consent, voluntary use, and anonymization techniques were among the most frequently fulfilled criteria. The least often implemented criteria were provisions of information about personal data breaches and data gathered from children. The balance between standards of data protection and public health benefits was achieved best by the COVIDSafe app and worst by the Alipay Health Code app.

Conclusions: Contact tracing apps with high levels of compliance with standards of data privacy tend to fulfill public health interests to a limited extent. Simultaneously, digital technologies with a lower level of data privacy protection allow for the collection of more data. Overall, this review shows that a consistent number of apps appear to comply with standards of data privacy, while their usefulness from a public health perspective can still be maximized.

(JMIR Mhealth Uhealth 2021;9(6):e23250) doi:10.2196/23250

KEYWORDS

COVID-19; contact tracing app; data accessibility; data privacy; mobile app; digital health; digital contact tracing

Introduction

The COVID-19 outbreak has shown how digital solutions can transform the health care system in an unprecedented manner. The rapid implementation of numerous innovative technologies to treat patients with COVID-19 has highlighted how much the human race can really benefit from data-driven transformation [1,2].

However, not all digital solutions have been launched in a contiguous manner. A specific example in this case includes mobile technologies. In principle, there are three areas in which mobile phone apps could aid in the fight against the COVID-19 pandemic: (1) self-diagnosis and facilitating treatment, (2) monitoring and enforcing the quarantine of infected persons, and (3) signaling if one is in close contact with an infected person [3].

The app used in the third area is called a *contact tracing app* and uses a smartphone to create a memory of any close contact with others for a significant amount of time. A warning is sent in case an infection is registered for anyone from the memory list, notifying the phone's owner to get tested and possibly self-quarantine.

Despite ongoing challenges to limit the spread of the virus, the launch of such digital solutions still faces several hurdles. Meanwhile, digital contact tracing apps have enormous potential, given the fact that there are more than 3.5 billion mobile phone users worldwide. However, there is significant resistance to allowing such technologies to interfere in the lives of individuals. More than one hundred nongovernmental organizations and civil rights associations have urged governments not to use the pandemic as an excuse to enter a new era of digital surveillance [4].

In an effort to overcome potential privacy challenges in the adoption of contact tracing technologies, the European Data Protection Board (EDPD) has published guidance for the use of location data and contact tracing tools intended to mitigate the impact of the COVID-19 pandemic [3]. Moreover, Article 9 (2) of the General Data Protection Regulation (GDPR) allows for access to special categories of data if the processing of such data is necessary for reasons of public interest in the public health sector, such as protection from serious cross-border threats to health. However, such a possibility consists of neither the full compression of privacy rights nor the GDPR itself [5].

Hence, existing guidelines and principles available at the European level should be considered as safe tools to guide developers of digital solutions, even in critical times, so as to ensure the adequate respect of individuals' rights. On the other hand, considering that the public health benefit of contact tracing apps depends directly on their widespread use, it is desirable to have governments strive to build trust among citizens with a high level of transparency. The pandemic showed how individual health is strictly connected to others' health, and that one of the most effective measures in the absence of a vaccine is the behavior of the individual, for that one person as much as for the whole community.

Therefore, while it is desirable to develop an app that guarantees an adequate level of privacy, it is equally desirable that citizens feel the need to use such an app as an act of social responsibility toward themselves and the community in general.

In this context, the objective of our study was to address the following two questions: (1) How do available contact tracing apps allow for data collection for public health benefits and comply with standards of data privacy? and (2) Can the balance between public health interests and the protection of personal data be established to ensure the broad use of digital technologies?

The success or failure of the adoption of contact tracing apps will have impacts beyond the ongoing fight against the COVID-19 pandemic. It will set a precedent for future opportunities and challenges in the integration of other digital solutions into clinical practice while ensuring the data privacy of its users. Therefore, we hope that our conclusions and recommendations can contribute to the debate regarding the adoption of digital technologies in the support of solving health issues, even beyond challenges related to the COVID-19 pandemic.

Methods

Systematic Review

We first performed a literature search in PubMed, MEDLINE, IEEE (Institute of Electrical and Electronics Engineers), and ACM (Association for Computing Machinery) Digital Library databases, covering the period between January 1 and August 31, 2020. The following key phrases were applied: "contact tracing," "contact detector," "contact mapping and COVID-19," "COVID-2019," "severe acute respiratory syndrome coronavirus 2," "2019-nCoV," and "SARS-CoV-2." The search terms are included in Multimedia Appendix 1. Only research articles written in English presenting a specific contact tracing app were included. Other publications, such as news articles, editorials, commentaries, reviews, or letters, were excluded. No geographical restriction was imposed. The selection and review of included publications were conducted independently by two reviewers, and discrepancies were resolved by consensus. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist was followed [6].

Supplementary Search

A supplementary search on the GitHub repository, as well as on governmental and app webpages, was then performed separately to collect additional information related to contact tracing apps that were identified in the systematic literature review. The following terms were screened: (1) overview, (2) frequently asked questions (FAQ), (3) data privacy, and (4) terms of use.

Assessment of the Fulfillment of Public Health Interests and Compliance With Data Privacy Guidelines

In order to address the first research question, two checklists were then developed to ensure a standardized approach toward the review of each technology. Both were sourced based on external reports in the fields of interest. The first checklist was constructed to review the key functions of mobile apps that define their fulfillment of public health interests. It was based on the Ada Lovelace Institute's report [7]. The second checklist was constructed to verify the compliance of contact tracing apps with selected data privacy standards. The following European guidelines were adopted in that respect: (1) the Privacy Code of Conduct for mobile health apps from the European Commission [8] and (2) the guidelines on the use of location data and contact tracing tools in the context of the COVID-19 outbreak from the EDPD [9]. The definition of each data privacy criterion that was used is available in Multimedia Appendix 2.

Balance Between Data Privacy and Public Health Interests

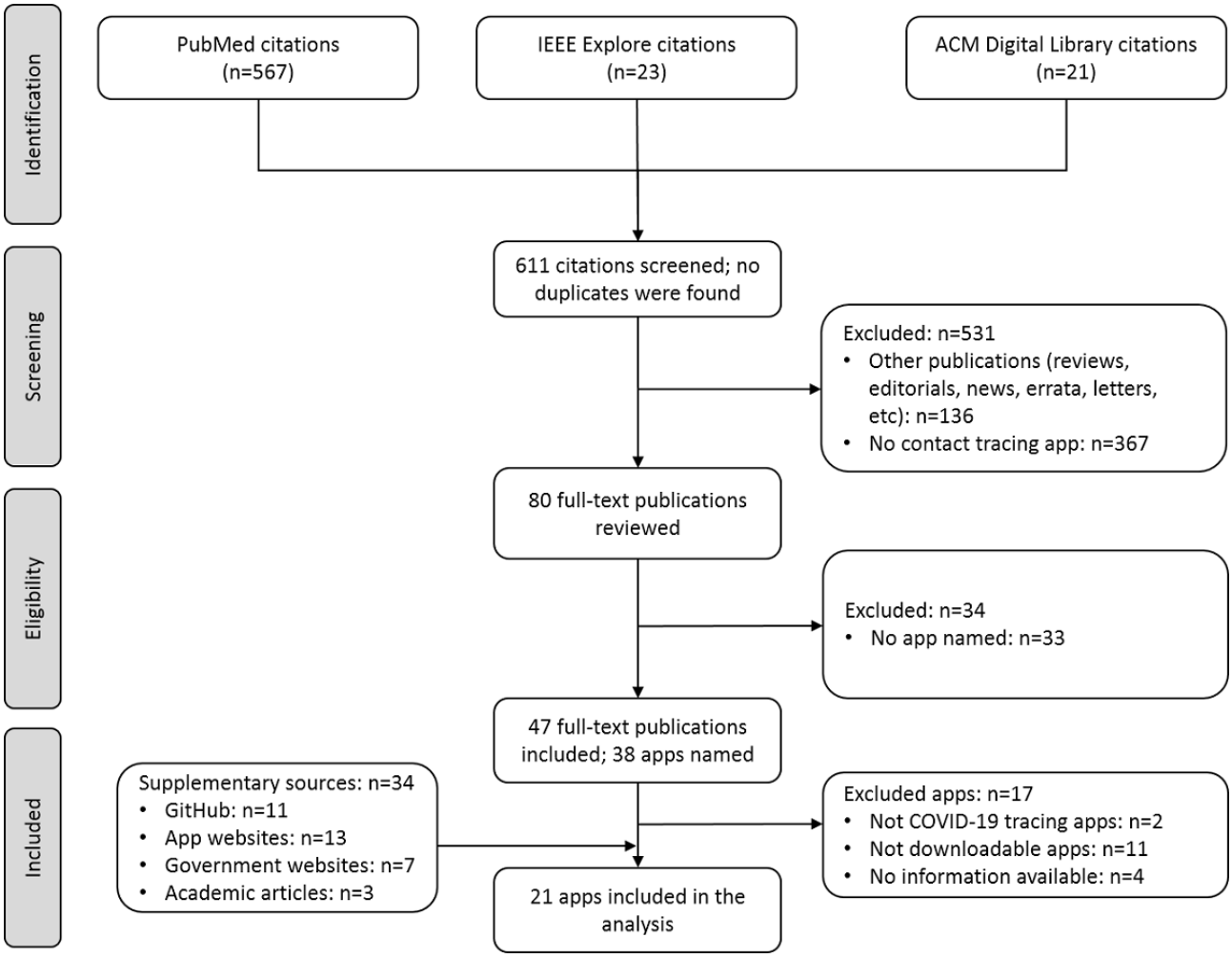
Finally, in order to address the second research question, a set of minimum requirements that simultaneously fulfill objectives of data privacy and public health was constructed. All reviewed contact tracing apps were ranked. The score of 1 or –1 was applied for each condition being met or not met, respectively. In the absence of information, a score of 0 was assigned. All requirements were considered as equally important. The ranking of all the reviewed technologies was constructed based on the number of scores granted.

Results

Systematic Review and Supplementary Research Eligibility

Overall, 611 unique records were identified across three databases, of which 531 articles did not meet the inclusion criteria (Figure 1). The abstract review of 80 full-text publications led to the exclusion of a further 33 articles (Multimedia Appendix 3). As the result of the full-text review of the remaining 47 publications, 38 different contact tracing apps were identified. Out of these 38 mobile apps, 17 were further excluded due to the unavailability of downloadable systems (n=11), a lack of sufficient information needed to perform the study (n=4), and not meeting criteria for contact tracing technology (n=2) (Figure 1). Further reasons for the exclusion of apps are provided in Multimedia Appendix 4 [10-42]. In the final set, 21 digital technologies were included. Among them were 18 (86%) and 3 (14%) apps that were used nationally and internationally, respectively. Additional information was collected from 34 supplementary references: government websites (n=7), app websites (n=13), academic articles (n=3), and GitHub (n=11) (Multimedia Appendix 5).

Figure 1. Flowchart of the article selection process for the literature search. ACM: Association for Computing Machinery; IEEE: Institute of Electrical and Electronics Engineers.



Assessment of the Apps' Fulfillment of Public Health Interests

In total, 11 criteria were defined to assess how digital technologies were able to fulfill public health interests. The three most frequently adopted types of functions in the group of reviewed contact tracing apps were (1) information about geographical coverage, (2) contact alerting, and (3) governmental responsibility. The two least frequently adopted functions were (1) medical and organizational support and (2) efficiency threshold (Multimedia Appendix 6 [10-27,43-56]). Out of the 21 apps, 19 (90%) received support from governments, but only 1 (5%) provided the additional threshold required to establish an app's efficiency. A confirmation of diagnosis with COVID-19 was required in 20 out of 21 (95%) cases, and 5 (24%) apps offered the possibility to register symptoms of COVID-19, such as fever, cough, or dyspnea. In 11 cases out of 21 (52%), the recommendation to self-isolate was issued. Other types of proximity data, such as recent travel, potentially related disorders, possible point of contact, and risk factors such as chronic disorders, were collected in half of apps as well (11/21, 52%) (Multimedia Appendix 6). Alipay Health Code was the only app that collected data about medical treatment; however, no details were found in that respect. Public health authorities and employers could access data stored in 10 (48%) and 1 (5%) contact tracing apps, respectively. Finally, there were two forms of support offered through contact tracing apps: a symptom checker and health care coordination (Multimedia Appendix 6).

Assessment of the Apps' Compliance With Data Privacy Guidelines

In total, 12 criteria were defined to verify to what extent digital technologies complied with data privacy guidelines. The three most frequently met conditions were (1) user consent, (2) voluntary basis, and (3) adoption of anonymization techniques (Multimedia Appendix 7 [10-27,43-56]). The provisions of information about (1) personal data breaches, (2) data gathered from children, and (3) confirmation of no data sharing to third parties were the least often implemented (Multimedia Appendix 7). As far as adequate security measures were concerned, 16 out of 21 (76%) contact tracing apps used a decentralized approach and 2 (10%) used a centralized approach. The remaining 3 (14%) apps did not provide such information. Moreover, 14 out of 21 (67%) apps published their code in the GitHub open source repository (Multimedia Appendix 7). Out of the 21 cases, 10 (48%) used anonymization techniques, with encryption methods being employed in 14 (67%) (Multimedia Appendix 7). Among other techniques, data aggregation, hashing algorithms, and asymmetrical unique keys were identified.

Furthermore, 15 of the 21 (74%) apps used pseudo-random (ie, frequently changing and ephemeral) identifiers as a method to ensure cybersecurity. This method suggests that the temporary ID of each device is created and modified periodically. Bluetooth appeared to be the most common method, which was adopted in 18 of the 21 (86%) apps, while 5 (24%) used GPS only, or in addition to Bluetooth. Regarding data retention, the average period of retention of proximity data was 3 weeks (Multimedia Appendix 7). Finally, 13 out of 21 (62%) apps used a mechanism to verify COVID-19–positive results, specifically verification through a code sent to a mobile phone or confirmed by health care professionals (Multimedia Appendix 7).

Balance Between Data Protection and Public Health Interests

The set of 10 and 6 minimum requirements related to data privacy and public health interests, respectively, were defined to assess the balance between these two domains (Multimedia Appendix 8). None of the analyzed apps managed to comply with all 16 conditions. COVIDSafe and SwissCovid met all 10 data privacy requirements and 5 out of 6 criteria for fulfillment of public health interests (15 points). The worst performance was achieved by Alipay Health Code (–3 points). Among the reviewed technologies, the standards of governmental accountability, anonymization, and encryption were the most frequently fulfilled. The establishment of an efficiency threshold and the adoption of rules against data breaches were the criteria with the lowest compliance.

Discussion

Principal Findings

The success of the digital revolution in the health care sector depends on access to a consistent volume of quality data that are useful to medical advancement, while ensuring an appropriate level of protection of personal data privacy. In our research, we attempted to systematically analyze the state of the art in the adoption of contact tracing apps in the fight against the COVID-19 pandemic, with the objective to verify to what extent the requirements concerning data protection and public health interests can be achieved simultaneously.

In order to develop the comprehensive list of criteria for our analysis, we searched the public domain for available guidelines in the fields of interest. We adopted a human-centered approach in that respect, which should be interpreted from two perspectives.

Firstly, when addressing the type of data being collected, we strived to choose the most comprehensive set of criteria that would ensure the maximization of public health benefit from a societal perspective. Therefore, we deliberately selected the Ada Lovelace Institute's report, which is based on the core values and the underlying mission of the institute (ie, to ensure that data and artificial intelligence work for people and society) [7].

Secondly, when addressing data protection, we focused on guidelines that were developed based on the most comprehensive set of data protection rules, such as the GDPR. From the European perspective, data protection is, in fact, considered a fundamental human right. Hence, we selected (1) the Privacy Code of Conduct on mobile health apps from the European Commission [8] and (2) the guidelines on the use of location data and contact tracing tools in the context of the COVID-19 outbreak from the EDPD. The GDPR is a regulation (ie, directly effective for member states without the need of further national laws) that concerns all types of processing of personal data and it applies to entities that process the personal data of European citizens. Hence, compliance with the GDPR is to be considered implicit for all European apps. The two selected legal instruments have a different level of authority, being a set of guidelines and a code of conduct, which means that developers can voluntarily commit to follow the rules. Hence, we decided to evaluate to what extent such rules—that we considered relevant in relation to contact tracing apps—have been followed by developers. Let us note that contact tracing apps can be considered mobile health technologies for two reasons: (1) information about COVID-19 positivity equates to health data and (2) there often are other consequences when the app sends a notification about contact with a positive case (eg, the receiver has a recommendation to self-isolate, whereas others need to contact health care providers). Hence, these consequences may be equivalent to the notion of *providing health advice*.

Not only did we develop the set of criteria for our analysis based on the chosen guidelines, but we also constructed a scorecard that simultaneously evaluated compliance with data protection and the level of fulfillment of public health benefits. It should be noted that although we did assign the same weight to all minimum requirements on the scorecard, we do value data protection as an essential condition due to the recognition of privacy as a human right and the urge to prevent a surveillance society. Hence, from our perspective, meeting the data protection criteria should be recommended as the mandatory base upon which to build digital health solutions and, consequently, public health benefits.

In sum, we reviewed 37 records about 21 contact tracing apps. In total, 12 and 11 criteria were used for the assessment of compliance with data privacy and fulfillment with public health interests, respectively. There are three key important findings worth highlighting.

Firstly, the majority of reviewed contact tracing apps were, to a greater extent, compliant with data privacy standards. As many as 18 out of 21 (86%) apps were designed to be used on a voluntary basis. Such an approach should be considered as optimal because it prevents any discrimination against individuals who are unable or unwilling to download the app. Moreover, the mandatory use of digital technology could lead to financial and technical burdens. Hence, jurisdictions with mandatory use of such apps are actually not compliant with the European standards. As demonstrated in Multimedia Appendix 7, BeAware Bahrain, Alipay Health Code, and Aarogya Setu were part of this category. A consistent number of apps used the decentralized approach (16/21, 76%), which is preferable to the centralized approach from a cybersecurity perspective. Additionally, Bluetooth technology with changing ephemeral

identifiers was employed in 15 out of 21 (71%) cases, which is preferable to GPS for privacy reasons, as Bluetooth is less intrusive. Still, there is room for improvement with respect to compliance with data privacy; in addition, the rules against data breaches were provided in only 3 out of 21 (14%) cases.

Secondly, as far as public health interests are concerned, there is still room for improvement among contact tracing apps. Only 8 out of 21 (38%) apps provided the definition of close contact, and governmental accountability was limited to 19 out of 21 (90%) cases. Additionally, 10 out of 21 (48%) apps reviewed solutions that were allowed for data accessibility for health care professionals. In that group, only 8 out of 10 (80%) apps provided details regarding techniques used for data aggregation, encryption, and/or anonymization. Meanwhile, there were just 13 out of 21 (62%) apps that used the mechanism of diagnosis verification, and the threshold for the population coverage needed for the digital solution to be effective was provided only once.

Thirdly, we observed the tendency that, with a high level of compliance with data privacy regulations, public health interests could be achieved to a limited extent, whereas a lower level of compliance with data privacy occurred across cases with greater potential for data collection. The COVIDSafe and SwissCovid apps met 15 out of 16 requirements. Both ensured that the minimum amount of data necessary for contact tracing was collected (ie, proximity data through Bluetooth). Contrary to that, Alipay Health Code ranked last in the scorecard as it processed a vast amount of data. Such a difference seems to be linked with different cultural, geographic, and political backgrounds. In principle, collecting more data could be useful from a public health perspective, for example, in order to closely monitor the contagion or to have a clearer and bigger picture of the spread of the disease from an epidemiological standpoint. Nonetheless, in relation to Alipay Health Code, no information was found about the actual use of the collected data for public health purposes. The lack of information available to users concerning the use of their data for further public health purposes would not, per se, infringe on the European principles as expressed in the GDPR. Indeed, Article 5.1 (b) of the GDPR considers further processing of collected data for public interest as compatible with the initial purpose of processing. Nonetheless, from a policy perspective, such ambiguity about the use of personal data might not be desirable in the European environment. In fact, not only might it disincentivize the use of apps due to the lack of transparency, but it could also have detrimental effects on society by undermining trust between citizens and governments.

Our findings need to be regarded with caution due to certain limitations of our study that must be acknowledged. Firstly, given the dynamic pandemic situation, we are aware that new developments in the field occur on a daily basis; as such, a number of new contact tracing apps were published after our systematic literature review was completed, or they were excluded due to limited information. On that note, it is worth mentioning that one of the most downloaded apps in the United States, namely HealthLynked COVID-19 Tracker [57,58], was not included in our study due to a lack of accessible information. Indeed, no sufficient information in relation to privacy and public health conditions was available in order to include the app in our tables. In fact, the website of the developer does not provide technical details nor a dedicated privacy policy for the tracking app; only the general privacy policy for other pre-existing HealthLynked apps was found. In our view, such disconnection between information provided to users and users' downloads of the app raises serious privacy concerns and does not constitute a good practice for future developers.

Secondly, let us highlight that our checklists were based on specific references, while there have been other relevant guidelines already developed in relation to COVID-19 contact tracing apps. An example in this case is the article *Ethical guidelines for COVID-19 tracing apps*, which was already published in *Nature* by a group of researchers from the University of Oxford after our study had been completed [10]. Such guidelines include not only principles of privacy but also relevant ethical issues, such as equality in the access to digital technology, which we do not discuss in our publication.

Thirdly, we deliberately omitted the assessment of effectiveness of reviewed contact tracing apps, as this was already published by other authors [59].

Finally, although we used a global approach to search for apps, our analysis was conducted from a European perspective. Non-European jurisdictions have diverse cultural and political backgrounds and, as such, may follow different sets of values [60]. Even before the COVID-19 outbreak, it had been routine to undergo temperature checks before visiting public places, maintain social distancing, and wear masks in some non-European countries [60]. Their technological apparatus is entirely different from the European one, as is the level of citizens' trust for the government. Accordingly, solutions adopted in such countries are not necessarily easily applicable in European settings and vice versa [60].

Conclusions

Despite the above limitations, we hope that our results and conclusions will inspire approaches on how to develop digital health solutions and ensure their broad adoption. Our literature review indicated that a consistent number of apps appear to be substantially compliant with the standards of data privacy, while the usefulness of contact tracing technologies from the public health perspective can still be maximized. Available surveys indicate that attitudes toward the use of contact tracing apps vary across different jurisdictions. According to a University of Oxford survey, the acceptance rate for contact tracing apps ranged from 67.5% to 85.5% in France, Germany, and Italy [61]. Conversely, a Pew Research Center study found that 60% of Americans believe that location tracking will not help limit the spread of COVID-19, and only 45% believe that such an app is allowed to track citizens who have had contact with an infected person [62]. Still, available estimates indicate that more than half of the population should use such an app in order to provide public health benefits [63]. Therefore, the question is how to ensure the broad adoption of contact tracing apps when their adoption happens on a voluntary basis, as was the case in 18 out of the 21 (86%) reviewed technologies. At the same time, only 1 of the 21 (5%) reviewed contact tracing apps provided information regarding the threshold for its efficiency in

combatting the COVID-19 pandemic. Consequently, the ultimate objective can be defined as the need to establish an evidence-based approach to the definition of the *broad use* of contact tracing apps and an educational campaign about their benefits. Indeed, mandatory use of digital health solutions would undermine the right to personal freedom that characterizes democratic societies. Another recent global survey on 7804 respondents from seven countries revealed that as many as 41% of the study population indicated data security as the number one barrier to adopting digital solutions. The adoption of the decentralized approach, which was the case for 16 out of 21 (76%) apps, may provide a better chance to obtain the required trust among end users [64].

Overall, a prima facie analysis suggests that the more intrusive the government is into an individual's privacy, the more that public health can benefit from the data. Moreover, it seems that a high level of privacy protection corresponds to an obstacle in terms of the use of data for public health. Nonetheless, we argue that the two interests can be perceived as not contradictory. Indeed, digital health solutions that protect individuals' privacy can be directed toward optimization of public health benefits. To achieve such a goal, the adoption of transparency policies that increase trust between public and private stakeholders should be encouraged. In fact, solid public confidence in digital solutions developed by governments that prioritize the protection of the rights of individuals can foster further data sharing for public health purposes, among other things.

Indeed, we believe that the key success factors in this matter are transparency and information campaigns targeting individuals, which can be achieved by working toward an awareness of citizens' responsibility and by providing relevant information on data protection and cybersecurity. Fostering citizen involvement in public matters such as public health can help to make every individual feel like a member of the state. Such awareness would stimulate individuals to share their data in a controlled and safe environment for the benefit of society.

Acknowledgments

During the conduct of this research, ZZ has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant 679681). This research was implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Tématerületi Kiválósági Program funding scheme (project No. TKP2020-NKA-02). MP has received funding from project No. 2019-1.3.1-KK-2019-00007, implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2019-1.3.1-KK funding scheme. FM received funding from the European Intellectual Property Institutes Network–Innovation Society (EIPIN IS) project funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant (agreement No. 721733).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms for the systematic review.
[DOCX File , 13 KB - mhealth_v9i6e23250_app1.docx]

Multimedia Appendix 2
Definitions of data privacy standard criteria.
[DOCX File , 25 KB - mhealth_v9i6e23250_app2.docx]

Multimedia Appendix 3
Reasons for exclusion of full-text articles.
[DOCX File , 16 KB - mhealth_v9i6e23250_app3.docx]

Multimedia Appendix 4
Reasons for exclusion of apps.
[DOCX File , 55 KB - mhealth_v9i6e23250_app4.docx]

Multimedia Appendix 5
Supplementary sources of information.
[DOCX File , 15 KB - mhealth_v9i6e23250_app5.docx]

Multimedia Appendix 6
Assessment of public health interests.
[DOCX File , 78 KB - mhealth_v9i6e23250_app6.docx]

Multimedia Appendix 7
Assessment of data privacy.
[DOCX File , 94 KB - mhealth_v9i6e23250_app7.docx]

Multimedia Appendix 8
Ranking of apps.
[XLSX File (Microsoft Excel File), 22 KB - mhealth_v9i6e23250_app8.xlsx]

References

1. Mońko M, Collins S. AI app, part-funded by EU, collects the sounds of COVID-19. European Research Council. 2020 Apr 06. URL: https://erc.europa.eu/news/ai-app-part-funded-eu-collects-sounds-covid-19?fbclid=IwAR2wQUYAI_8nsBqI2_YJgH2N1XaRNfQUq5VeECu891-6Mqd8iexJJ9Eq2XI [accessed 2020-05-08]
2. Siwicki B. Americans' perceptions of telehealth in the COVID-19 era. Healthcare IT News. 2020 Apr 03. URL: https://www.healthcareitnews.com/news/survey-americans-perceptions-telehealth-covid-19-era [accessed 2020-05-08]
3. Commission recommendation (EU) 2020/518 of 8 April 2020 on a common Union toolbox for the use of technology and data to combat and exit from the COVID-19 crisis, in particular concerning mobile applications and the use of anonymized mobility data. European Union. 2020 Apr 08. URL: https://op.europa.eu/pl/publication-detail/-/publication/1e8b1520-7e0c-11ea-aea8-01aa75ed71a1/language-en [accessed 2020-05-05]
4. Digital surveillance to fight COVID-19 can only be justified if it respects human rights. Amnesty International. 2020 Apr 02. URL: https://www.amnesty.org/en/latest/news/2020/04/covid19-digital-surveillance-ngo/ [accessed 2020-05-10]
5. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Eur-Lex. 2016 Apr 27. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2020-05-05]
6. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
7. Exit through the App Store? Rapid evidence review. Ada Lovelace Institute. London, UK: Ada Lovelace Institute; 2020 Apr 20. URL: https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-Rapid-Evidence-Review-Exit-through-the-App-Store-April-2020-2.pdf [accessed 2020-05-05]
8. Privacy code of conduct on mobile health apps. European Commission. URL: https://ec.europa.eu/digital-single-market/en/privacy-code-conduct-mobile-health-apps [accessed 2020-05-04]
9. Guidelines 04/2020 on the Use of Location Data and Contact Tracing Tools in the Context of the COVID-19 Outbreak. Brussels, Belgium: European Data Protection Board; 2020 Apr 21. URL: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf [accessed 2020-05-15]
10. Morley J, Cowls J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. Nature 2020 Jun;582(7810):29-31. [doi: 10.1038/d41586-020-01578-0] [Medline: 32467596]
11. Mbunge E. Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. Diabetes Metab Syndr 2020;14(6):1631-1636 [FREE Full text] [doi: 10.1016/j.dsx.2020.08.029] [Medline: 32892060]
12. Maghdid HS, Ghafoor KZ. A smartphone enabled approach to manage COVID-19 lockdown and economic crisis. SN Comput Sci 2020;1(5):271 [FREE Full text] [doi: 10.1007/s42979-020-00290-0] [Medline: 33063052]
13. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]
14. Garg L, Chukwu E, Nasser N, Chakraborty C, Garg G. Anonymity preserving IoT-based COVID-19 and other infectious disease contact tracing model. IEEE Access 2020;8:159402-159414. [doi: 10.1109/access.2020.3020513]
15. Bradford L, Aboy M, Liddell K. COVID-19 contact tracing apps: A stress test for privacy, the GDPR, and data protection regimes. J Law Biosci 2020;7(1):lsaa034 [FREE Full text] [doi: 10.1093/jlb/lsaa034] [Medline: 32728470]
16. Kricka LJ, Polevikov S, Park JY, Fortina P, Bernardini S, Satchkov D, et al. Artificial intelligence-powered search tools and resources in the fight against COVID-19. EJIFCC 2020 Jun;31(2):106-116 [FREE Full text] [Medline: 32549878]
17. Lin L, Hou Z. Combat COVID-19 with artificial intelligence and big data. J Travel Med 2020 Aug 20;27(5):taaa080 [FREE Full text] [doi: 10.1093/jtm/taaa080] [Medline: 32437541]
18. Schneble CO, Elger BS, Martin Shaw D. Data protection during the coronavirus crisis. EMBO Rep 2020 Sep 03;21(9):e51362 [FREE Full text] [doi: 10.15252/embr.202051362] [Medline: 32783305]
19. Leins K, Culnane C, Rubinstein BI. Tracking, tracing, trust: Contemplating mitigating the impact of COVID-19 with technological interventions. Med J Aust 2020 Jul;213(1):6-8.e1 [FREE Full text] [doi: 10.5694/mja2.50669] [Medline: 32548879]

20. Wise J. Covid-19: UK drops its own contact tracing app to switch to Apple and Google model. BMJ 2020 Jun 19;369:m2472. [doi: 10.1136/bmj.m2472] [Medline: 32561511]
21. Park YJ, Cho SY, Lee J, Lee I, Park W, Jeong S, et al. Development and utilization of a rapid and accurate epidemic investigation support system for COVID-19. Osong Public Health Res Perspect 2020 Jun;11(3):118-127 [FREE Full text] [doi: 10.24171/j.phrp.2020.11.3.06] [Medline: 32528817]
22. Nuzzo A, Tan CO, Raskar R, DeSimone DC, Kapa S, Gupta R. Universal shelter-in-place versus advanced automated contact tracing and targeted isolation: A case for 21st-century technologies for SARS-CoV-2 and future pandemics. Mayo Clin Proc 2020 Sep;95(9):1898-1905 [FREE Full text] [doi: 10.1016/j.mayocp.2020.06.027] [Medline: 32861334]
23. Almeida BDA, Doneda D, Ichihara MY, Barral-Netto M, Matta GC, Rabello ET, et al. Personal data usage and privacy considerations in the COVID-19 global pandemic. Cien Saude Colet 2020 Jun;25(suppl 1):2487-2492 [FREE Full text] [doi: 10.1590/1413-81232020256.1.11792020] [Medline: 32520293]
24. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-peer contact tracing: Development of a privacy-preserving smartphone app. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
25. Hernandez-Orallo E, Manzoni P, Calafate CT, Cano J. Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: The case of COVID-19. IEEE Access 2020;8:99083-99097. [doi: 10.1109/access.2020.2998042]
26. Rowe F. Contact tracing apps and values dilemmas: A privacy paradox in a neo-liberal world. Int J Inf Manage 2020 Dec;55:102178 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102178] [Medline: 32836636]
27. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: Taxonomy, ethical challenges, and navigation aid. Lancet 2020 Aug;2(8):e425-e434.
28. Wilmink G, Summer I, Marsyla D, Sukhu S, Grote J, Zobel G, et al. Real-time digital contact tracing: Development of a system to control COVID-19 outbreaks in nursing homes and long-term care facilities. JMIR Public Health Surveill 2020 Aug 25;6(3):e20828 [FREE Full text] [doi: 10.2196/20828] [Medline: 32745013]
29. Vokinger KN, Nittas V, Witt CM, Fabrikant SI, von Wyl V. Digital health and the COVID-19 epidemic: An assessment framework for apps from an epidemiological and legal perspective. Swiss Med Wkly 2020 May 04;150:w20282 [FREE Full text] [doi: 10.4414/smw.2020.20282] [Medline: 32418194]
30. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
31. Cohen IG, Gostin LO, Weitzner DJ. Digital smartphone tracking for COVID-19: Public health and civil liberties in tension. JAMA 2020 Jun 16;323(23):2371-2372. [doi: 10.1001/jama.2020.8570] [Medline: 32459289]
32. Sharon T. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. Ethics Inf Technol 2020 Jul 18:1-13 [FREE Full text] [doi: 10.1007/s10676-020-09547-x] [Medline: 32837287]
33. Martinez-Martin N, Wieten S, Magnus D, Cho MK. Digital contact tracing, privacy, and public health. Hastings Cent Rep 2020 May;50(3):43-46 [FREE Full text] [doi: 10.1002/hast.1131] [Medline: 32596893]
34. White L, van Basshuysen P. How to overcome lockdown: Selective isolation versus contact tracing. J Med Ethics 2020 Nov;46(11):724-725 [FREE Full text] [doi: 10.1136/medethics-2020-106680] [Medline: 32817409]
35. Baraniuk C. Covid-19 contact tracing: A briefing. BMJ 2020 May 13;369:m1859. [doi: 10.1136/bmj.m1859] [Medline: 32404304]
36. Hegde A, Masthi R. Digital contact tracing in the COVID-19 pandemic: A tool far from reality. Digit Health 2020;6:1-3 [FREE Full text] [doi: 10.1177/2055207620946193] [Medline: 32939289]
37. Labs J, Terry S. Privacy in the coronavirus era. Genet Test Mol Biomarkers 2020 Sep;24(9):535-536. [doi: 10.1089/gtmb.2020.29055.sjt] [Medline: 32677850]
38. Nguyen CT, Saputra YM, Van Huynh N, Nguyen N, Khoa TV, Tuan BM, et al. A comprehensive survey of enabling and emerging technologies for social distancing—Part II: Emerging technologies and open issues. IEEE Access 2020;8:154209-154236. [doi: 10.1109/access.2020.3018124]
39. Nguyen CT, Saputra YM, Huynh NV, Nguyen N, Khoa TV, Tuan BM, et al. A comprehensive survey of enabling and emerging technologies for social distancing—Part I: Fundamentals and enabling technologies. IEEE Access 2020;8:153479-153507. [doi: 10.1109/access.2020.3018140]
40. Kleinman RA, Merkel C. Digital contact tracing for COVID-19. CMAJ 2020 Jun 15;192(24):E653-E656 [FREE Full text] [doi: 10.1503/cmaj.200922] [Medline: 32461324]
41. Kaspar K. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: Quantitative survey study. J Med Internet Res 2020 Aug 27;22(8):e21613 [FREE Full text] [doi: 10.2196/21613] [Medline: 32759100]
42. Ho HJ, Zhang ZX, Huang Z, Aung AH, Lim W, Chow A. Use of a real-time locating system for contact tracing of health care workers during the COVID-19 pandemic at an infectious disease center in Singapore: Validation study. J Med Internet Res 2020 May 26;22(5):e19437 [FREE Full text] [doi: 10.2196/19437] [Medline: 32412416]
43. Gupta R, Bedi M, Goyal P, Wadhera S, Verma V. Analysis of COVID-19 tracking tool in India. Digit Gov 2020 Dec 03;1(4):1-8. [doi: 10.1145/3416088]

44. Baumgart DC. Digital advantage in the COVID-19 response: Perspective from Canada's largest integrated digitalized healthcare system. NPJ Digit Med 2020;3:114 [FREE Full text] [doi: 10.1038/s41746-020-00326-y] [Medline: 32923691]
45. Ye Q, Zhou J, Wu H. Using information technology to manage the COVID-19 pandemic: Development of a technical framework based on practical experience in China. JMIR Med Inform 2020 Jun 08;8(6):e19515 [FREE Full text] [doi: 10.2196/19515] [Medline: 32479411]
46. Lucivero F, Hallowell N, Johnson S, Prainsack B, Samuel G, Sharon T. COVID-19 and contact tracing apps: Ethical challenges for a social experiment on a global scale. J Bioeth Inq 2020 Dec;17(4):835-839 [FREE Full text] [doi: 10.1007/s11673-020-10016-9] [Medline: 32840842]
47. Owusu PN. Digital technology applications for contact tracing: The new promise for COVID-19 and beyond? Glob Health Res Policy 2020;5:36 [FREE Full text] [doi: 10.1186/s41256-020-00164-1] [Medline: 32766453]
48. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
49. Ekong I, Chukwu E, Chukwu M. COVID-19 mobile positioning data contact tracing and patient privacy regulations: Exploratory search of global response strategies and the use of digital tools in Nigeria. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e19139 [FREE Full text] [doi: 10.2196/19139] [Medline: 32310817]
50. Braun P, Haffner S, Woodcock BG. COVID-19 pandemic predictions using the modified Bateman SIZ model and observational data for Heidelberg, Germany: Effect of vaccination with a SARS-CoV-2 vaccine, coronavirus testing and application of the Corona-Warn-App. Int J Clin Pharmacol Ther 2020 Aug;58(8):417-425. [doi: 10.5414/CP203846] [Medline: 32646540]
51. Couch DL, Robinson P, Komesaroff PA. COVID-19-extending surveillance and the panopticon. J Bioeth Inq 2020 Dec;17(4):809-814 [FREE Full text] [doi: 10.1007/s11673-020-10036-5] [Medline: 32840859]
52. Currie DJ, Peng CQ, Lyle DM, Jameson BA, Frommer MS. Stemming the flow: How much can the Australian smartphone app help to control COVID-19? Public Health Res Pract 2020 Jun 30;30(2):1-11 [FREE Full text] [doi: 10.17061/phrp3022009] [Medline: 32601652]
53. Iacobucci G. Covid-19: Is local contact tracing the answer? BMJ 2020 Aug 17;370:m3248. [doi: 10.1136/bmj.m3248] [Medline: 32816820]
54. Ryan BJ, Coppola D, Williams J, Swienton R. COVID-19 contact tracing solutions for mass gatherings. Disaster Med Public Health Prep 2020 Jul 14:1-7 [FREE Full text] [doi: 10.1017/dmp.2020.241] [Medline: 32660677]
55. Wallentin G, Kaziyeva D, Reibersdorfer-Adelsberger E. COVID-19 intervention scenarios for a long-term disease management. Int J Health Policy Manag 2020 Dec 01;9(12):508-516 [FREE Full text] [doi: 10.34172/ijhpm.2020.130] [Medline: 32729281]
56. 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] [Medline: 32340107]
57. 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] [Medline: 32609622]
58. Brewster T. Google bans coronavirus infection trackers... but not before they get 400,000 downloads. Forbes. 2020 Mar 24. URL: https://www.forbes.com/sites/thomasbrewster/2020/03/24/google-bans-coronavirus-apps-but-after-400000-downloads/?sh=3c095d9234c0 [accessed 2020-05-10]
59. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: A systematic review to inform the control of COVID-19. Lancet 2020 Nov;2(11):e607-e621.
60. Saetta B. Il caso Singapore, il sistema investigativo (non solo) tecnologico per contrastare le epidemie, e la privacy. Valigia Blu. 2020 Mar 28. URL: https://www.valigiablu.it/approfondimenti/il-caso-singapore-il-sistema-investigativo-non-solo-tecnologico-per-contrastare-le-epidemie-e-la-privacy/ [accessed 2020-05-15]
61. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: Cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
62. Anderson M, Auxier B. Most Americans don't think cellphone tracking will help limit COVID-19, are divided on whether it's acceptable. Pew Research Center. Washington, DC: Pew Research Center; 2020 Apr 16. URL: https://www.pewresearch.org/fact-tank/2020/04/16/most-americans-dont-think-cellphone-tracking-will-help-limit-covid-19-are-divided-on-whether-its-acceptable/ [accessed 2020-05-15]
63. Britton T, Ball F, Trapman P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 2020 Aug 14;369(6505):846-849 [FREE Full text] [doi: 10.1126/science.abc6810] [Medline: 32576668]

64. Digital is transforming health, so why is consumer adoption stalling? Accenture 2020 Digital Health Customer Survey US results. Accenture. 2020. URL: https://www.accenture.com/_acnmedia/PDF-118/Accenture-2020-digital-health-consumer-survey.pdf [accessed 2020-05-15]

Abbreviations

ACM: Association for Computing Machinery
EDPD: European Data Protection Board
EIPIN IS: European Intellectual Property Institutes Network–Innovation Society
FAQ: frequently asked questions
GDPR: General Data Protection Regulation
IEEE: Institute of Electrical and Electronics Engineers
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 05.08.20; peer-reviewed by J D'Arcey, A Teles, O El-Gayar; comments to author 07.10.20; revised version received 30.10.20; accepted 22.02.21; published 10.06.21.

Please cite as:
Kolasa K, Mazzi F, Leszczuk-Czubkowska E, Zrubka Z, Péntek M
State of the Art in Adoption of Contact Tracing Apps and Recommendations Regarding Privacy Protection and Public Health: Systematic Review
JMIR Mhealth Uhealth 2021;9(6):e23250
URL: https://mhealth.jmir.org/2021/6/e23250
doi:10.2196/23250
PMID:34033581

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org